

| Result No. | Score | Query | | ID | Description |
|------------|-------|-------|-----------|----|-------------|
| | | Match | Length DB | | |
| 1 | 118.5 | 8.2 | 233 | 1 | TNFA_CANFA |
| 2 | 110.5 | 7.6 | 235 | 1 | TNFA_RAT |
| 3 | 109.5 | 7.5 | 233 | 1 | TNFA_FELCA |
| 4 | 109.5 | 7.5 | 233 | 1 | TNFA_MOUSE |
| 5 | 108 | 7.4 | 235 | 1 | TNFA_BOVIN |
| 6 | 107.5 | 7.4 | 235 | 1 | TNFA_PERLE |
| 7 | 107 | 7.4 | 193 | 1 | TNFA_CAPHI |
| 8 | 107 | 7.4 | 232 | 1 | TNFA_PIG |
| 9 | 102.5 | 7.1 | 234 | 1 | TNFA_SHEEP |
| 10 | 101.5 | 7.0 | 229 | 1 | TNFA_CEREL |
| 11 | 100.5 | 6.9 | 205 | 1 | TNFB_HUMAN |
| 12 | 99 | 6.8 | 234 | 1 | TNFA_CAVPO |
| 13 | 98.5 | 6.8 | 233 | 1 | TNFA_HUMAN |
| 14 | 97.5 | 6.7 | 281 | 1 | FASL_HUMAN |
| 15 | 97 | 6.7 | 281 | 1 | TRAI_HUMAN |
| 16 | 95 | 6.5 | 233 | 1 | TNFA_TRIVU |
| 17 | 93.5 | 6.4 | 651 | 1 | E2BD_YEAST |
| 18 | 93.5 | 6.4 | 235 | 1 | TNFB_RABIT |
| 19 | 90.5 | 6.2 | 204 | 1 | TNFB_PIG |
| 20 | 89.5 | 6.2 | 234 | 1 | TNFA_HORSE |
| 21 | 89.5 | 6.2 | 233 | 1 | TNFA_PAPSP |
| 22 | 87.5 | 6.0 | 993 | 1 | TSH_DROME |
| 23 | 85.5 | 5.9 | 233 | 1 | TNFA_MACFA |
| 24 | 85.5 | 5.9 | 233 | 1 | TNFA_MACMU |
| 25 | 85.5 | 5.9 | 204 | 1 | TNFB_BOVIN |
| 26 | 85.5 | 5.9 | 358 | 1 | VE2_HPV51 |
| 27 | 85.5 | 5.9 | 351 | 1 | YIIG_ECOLI |
| 28 | 84.5 | 5.8 | 291 | 1 | TRAI_MOUSE |
| 29 | 84 | 5.8 | 197 | 1 | TNFB_RABIT |
| 30 | 84 | 5.8 | 295 | 1 | YFLF_CAUCR |
| 31 | 82.5 | 5.7 | 373 | 1 | DHA_MYCTU |
| 32 | 82.5 | 5.7 | 680 | 1 | NASP_RABIT |
| 33 | 82.5 | 5.7 | 478 | 1 | OCY2_PIG |
| 34 | 82.5 | 5.7 | 593 | 1 | STX2_CANFA |
| 35 | 81.5 | 5.6 | 279 | 1 | FASL_MOUSE |
| 36 | 81 | 5.6 | 278 | 1 | FASL_RAT |
| 37 | 81 | 5.6 | 570 | 1 | HAQ_NITEUP |
| 38 | 81 | 5.6 | 769 | 1 | VPI_BPpH6 |
| 39 | 80.5 | 5.5 | 2067 | 1 | BIMB_EMENI |
| 40 | 79.5 | 5.5 | 1150 | 1 | 2ACA_HUMAN |
| 41 | 79.5 | 5.5 | 889 | 1 | COPP_YEAST |
| 42 | 79.5 | 5.5 | 2032 | 1 | CTOG_HUMAN |
| 43 | 79 | 5.4 | 561 | 1 | BCHD_RHOCA |

 • •

 • •

 • •
 • •

| Category | Frequency | Percentage | Mean | Standard Deviation | Minimum | Maximum |
|----------|-----------|------------|-------|--------------------|---------|---------|
| 1 | 1 | 100.00% | 1.00 | 0.00 | 1 | 1 |
| 2 | 1 | 100.00% | 2.00 | 0.00 | 2 | 2 |
| 3 | 1 | 100.00% | 3.00 | 0.00 | 3 | 3 |
| 4 | 1 | 100.00% | 4.00 | 0.00 | 4 | 4 |
| 5 | 1 | 100.00% | 5.00 | 0.00 | 5 | 5 |
| 6 | 1 | 100.00% | 6.00 | 0.00 | 6 | 6 |
| 7 | 1 | 100.00% | 7.00 | 0.00 | 7 | 7 |
| 8 | 1 | 100.00% | 8.00 | 0.00 | 8 | 8 |
| 9 | 1 | 100.00% | 9.00 | 0.00 | 9 | 9 |
| 10 | 1 | 100.00% | 10.00 | 0.00 | 10 | 10 |
| 11 | 1 | 100.00% | 11.00 | 0.00 | 11 | 11 |
| 12 | 1 | 100.00% | 12.00 | 0.00 | 12 | 12 |
| 13 | 1 | 100.00% | 13.00 | 0.00 | 13 | 13 |
| 14 | 1 | 100.00% | 14.00 | 0.00 | 14 | 14 |
| 15 | 1 | 100.00% | 15.00 | 0.00 | 15 | 15 |
| 16 | 1 | 100.00% | 16.00 | 0.00 | 16 | 16 |
| 17 | 1 | 100.00% | 17.00 | 0.00 | 17 | 17 |
| 18 | 1 | 100.00% | 18.00 | 0.00 | 18 | 18 |
| 19 | 1 | 100.00% | 19.00 | 0.00 | 19 | 19 |
| 20 | 1 | 100.00% | 20.00 | 0.00 | 20 | 20 |
| 21 | 1 | 100.00% | 21.00 | 0.00 | 21 | 21 |
| 22 | 1 | 100.00% | 22.00 | 0.00 | 22 | 22 |
| 23 | 1 | 100.00% | 23.00 | 0.00 | 23 | 23 |
| 24 | 1 | 100.00% | 24.00 | 0.00 | 24 | 24 |
| 25 | 1 | 100.00% | 25.00 | 0.00 | 25 | 25 |
| 26 | 1 | 100.00% | 26.00 | 0.00 | 26 | 26 |
| 27 | 1 | 100.00% | 27.00 | 0.00 | 27 | 27 |
| 28 | 1 | 100.00% | 28.00 | 0.00 | 28 | 28 |
| 29 | 1 | 100.00% | 29.00 | 0.00 | 29 | 29 |
| 30 | 1 | 100.00% | 30.00 | 0.00 | 30 | 30 |
| 31 | 1 | 100.00% | 31.00 | 0.00 | 31 | 31 |
| 32 | 1 | 100.00% | 32.00 | 0.00 | 32 | 32 |
| 33 | 1 | 100.00% | 33.00 | 0.00 | 33 | 33 |
| 34 | 1 | 100.00% | 34.00 | 0.00 | 34 | 34 |
| 35 | 1 | 100.00% | 35.00 | 0.00 | 35 | 35 |
| 36 | 1 | 100.00% | 36.00 | 0.00 | 36 | 36 |
| 37 | 1 | 100.00% | 37.00 | 0.00 | 37 | 37 |
| 38 | 1 | 100.00% | 38.00 | 0.00 | 38 | 38 |
| 39 | 1 | 100.00% | 39.00 | 0.00 | 39 | 39 |
| 40 | 1 | 100.00% | 40.00 | 0.00 | 40 | 40 |
| 41 | 1 | 100.00% | 41.00 | 0.00 | 41 | 41 |
| 42 | 1 | 100.00% | 42.00 | 0.00 | 42 | 42 |
| 43 | 1 | 100.00% | 43.00 | 0.00 | 43 | 43 |
| 44 | 1 | 100.00% | 44.00 | 0.00 | 44 | 44 |
| 45 | 1 | 100.00% | 45.00 | 0.00 | 45 | 45 |
| 46 | 1 | 100.00% | 46.00 | 0.00 | 46 | 46 |
| 47 | 1 | 100.00% | 47.00 | 0.00 | 47 | 47 |
| 48 | 1 | 100.00% | 48.00 | 0.00 | 48 | 48 |
| 49 | 1 | 100.00% | 49.00 | 0.00 | 49 | 49 |
| 50 | 1 | 100.00% | 50.00 | 0.00 | 50 | 50 |
| 51 | 1 | 100.00% | 51.00 | 0.00 | 51 | 51 |
| 52 | 1 | 100.00% | 52.00 | 0.00 | 52 | 52 |
| 53 | 1 | 100.00% | 53.00 | 0.00 | 53 | 53 |
| 54 | 1 | 100.00% | 54.00 | 0.00 | 54 | 54 |
| 55 | 1 | 100.00% | 55.00 | 0.00 | 55 | 55 |
| 56 | 1 | 100.00% | 56.00 | 0.00 | 56 | 56 |
| 57 | 1 | 100.00% | 57.00 | 0.00 | 57 | 57 |
| 58 | | | | | | |

QY 60 CLTVSFYVAALQGLDLASLRAELQGHAEKLPAGAGAPKAGLEEAPAVTAGLKIFEPPA 119

```

Db 32 CLSLFSLVAGATTLCFLHFGVIGPQREELP-----NGQLISPLA 74
QY 120 PGGNSSONSRNK---RANQCPETVTQDCLQLIADSETPIQKGYTFVPWLLSKRGS 176
Db 75 QTVKSSSRTPSDKPVAVVANPE-----AEGQ-----LQWL--SRAN 110
QY 177 AL-----EKENILVKETGYFFIYGOVLVTDKTYAMGHILQKVKHVFQ---DELSLV 227
Db 111 ALLANGVELTDNLQVLPSDGLYLIYQVLFKGGCGSTHVLTHTSRFAVSTQTKVNL 170
QY 228 TLFR--CIONMPETLPNNSCYS-----AGIAKLEEGDELQLAIPRENAQISLDGDTFFGA 281
Db 171 SAIKSPCQRETPREGTEAKPWYEDILYLGGVFQLEKGDRLSAEINLPNVLDPFAESGQVYFGI 230
QY 282 LKL 284
Db 231 IAL 233

RESULT 2
TNFA_RAT
ID TNFA_RAT STANDARD; PRT; 235 AA.
AC P16599;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=TESTIS;
RX MEDLINE; 94040766.
RA KWON J., CHUNG I.Y., BENVENISTE E.N.;
RL GENE 132:227-236(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA SHIRAI T., SHIMIZU N., HORIGUCHI S., ITO H.;
RL AGRIC. BIOL. CHEM. 53:1733-1736(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE; 92329007.
RA ESTLER H.C., GREWE M., GAUSSLING R., PAVLOVIC M., DECKER K.;
RL BIOL. CHEM. HOPPE-SEYLER 373:271-281(1992).
RN [4]
RP SEQUENCE FROM N.A.
RA DECKER K.F.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]
RP SEQUENCE OF 1-231 FROM N.A.
RC TISSUE=TAIL;
RA KIRISITS M.J., VARDIMON D., KUNZ H.W., GILL T.J. III;
RL SUBMITTED (JUN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -!- SUBUNIT: HOMOTRIMER.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL; X66539; G395370; -.
DR EMBL; L00981; G205254; -.

```

```

DR EMBL; D00475; G220921; -.
DR EMBL; AJ002278; E1154446; -.
DR EMBL; L19123; G310232; -.
DR PIR; JU0029; JU0029.
DR PIR; S21674; S21674.
DR PIR; JN0868; JN0868.
DR HSSP; P01375; ITNF.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS50049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 79
FT CHAIN 80 235 TUMOR NECROSIS FACTOR.
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DISULFID 148 179 BY SIMILARITY.
FT CARBOHYD 86 86 POTENTIAL.
FT CONFLICT 39 39 L -> P (IN REF. 3 AND 4).
FT CONFLICT 163 163 I -> T (IN REF. 3 AND 4).
FT CONFLICT 202 202 F -> S (IN REF. 3 AND 4).
SQ SEQUENCE 235 AA; 25806 MW; CIDF18B7 CRC32;

Query Match 7.6%; Score 110.5; DB 1; Length 235;
Best Local Similarity 22.2%; Pred. No. 0.017;
Matches 54; Conservative 45; Mismatches 87; Indels 57; Gaps 11;

QY 60 CLTVSVFYOVAALQGLDLSRAELQG-HHAEKLPAGAGAPKAGLEAPAVTAGLKTFEPP 118
Db 32 CLSLFSLVAGATTLCFLHFGVIGPQREELPNG-----LPLISSMAQTTLR----- 82
QY 119 AFGEGNSSONSRKRAVQPEETVTQDCLQLIADSETPIQKGYTFVPWLLSKRGSAL 178
Db 82 -----SSONSSDKPVAVHVAHQAEQLEWLSQRANALLANG-----M 120
QY 179 EKENILVKETGYFFIYGOVLVTDK---TYAMGHILQKVKHVFQDELSTLTLFR--C 232
Db 121 DLKDNQVVPADGLYLIYQVLFKGGCGPDYVLLHTVTSRFAIS-YQEKVSLLSAKSPC 179
QY 233 IONMPETLP-----NNSCYSAGIAKLEEGDELQLAIPRENAQISLDG--DVT-----FFGA 281
Db 180 PKDTPGAEKLPWEPYMLGGVQLEKGDLL-----SAEVLNPKYLDITESGQVYFVG 232
QY 282 LKL 284
Db 233 IAL 235

RESULT 3
TNFA_FELCA
ID TNFA_FELCA STANDARD; PRT; 233 AA.
AC P19101;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS FELIS SILVESTRIS CATUS (CAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; CARNIVORA.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BLOOD;
RX MEDLINE; 91016860.
RA MCGRAW R.A., COFFEE B.W., OTTO C.M., DREWS R.T., RAWLINGS C.A.;
RL NUCLEIC ACIDS RES. 18:5563-5563(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=BONE MARROW;
RA DANIEL S.L., BRENNER C.A., LEGENDRE A.M., SOLOMAN A., ROUSE B.T.;
RL SUBMITTED (XXX-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION

```



```
RA CREW M.D., FILIPOWSKI M.E.;
RL IMMUNOGENETICS 35:351-353(1992).
CC -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -!- SUBUNIT: HOMOTRIMER.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL; M59233; G202507; -.
DR HSSP; P01375; ITNF.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS00049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 79 BY SIMILARITY.
FT CHAIN 80 235 TUMOR NECROSIS FACTOR.
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DISULFID 148 179 BY SIMILARITY.
FT CARBOHYD 86 86 POTENTIAL.
SQ SEQUENCE 235 AA; 25822 MW; 0D9AF4E8 CRC32;

Query Match 7.4%; Score 107.5; DB 1; Length 235;
Best Local Similarity 22.0%; Pred. No. 0.031;
Matches 54; Conservative 46; Mismatches 84; Indels 61; Gaps 14;

QY 60 CLTVVFTQVAALQGLASLRAELAG-HIAEKLPGAGAPKAGLEEAPAVTAGLKIFEP 118
DQ :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: |||
DB 32 CLSLFSLVLVAGATTLCFLNFGVGPQREKFP--NNLPIIG---SMAQTTLR----- 82
QY 119 APGEGNSSNSNRKAVQGPETVTQDCLQIADSETPTIQKGSYTFVFWLLSFKRGSAL 178
DB :||||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: |||
DB 82 -----SSSONSSDK-----PVAHVVAHQVDEQLWLSRGNALL-----ANGM 120
QY 179 EEKENKILVKETGYFFIYGVLYTDK---TYA-MGLHIOKKVHVFGDELSLVTIFRCIQ 234
DB :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: |||
DB 121 DLKDNQVLPADGLVLYSVQLFKQGCSSYVLLHTVSRFAVS-YEDKVNLSAIK--S 177
QY 235 NNPTLNNSS-----CYSAGIAKLEGEDEL--QLAIPR-----ENAIQLDGDGVTF 279
DB :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: |||
DB 178 PCPKETPEGSELKPWYEPYILGVGFQLEKGRDLSAEVNLPKYLDFAESQV-----YF 230
QY 280 GALKL 284
DB :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: |||
DB 231 GVIAL 235

RESULT 7
TNFA_CAPHI
ID TNFA_CAPHI STANDARD: PRT: 193 AA.
AC P13296;
DT 01-JAN-1990 (REL. 13, CREATED)
DT 01-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN) (FRAGMENT).
GN TNFA.
OS CAPRA HIRCUS (GOAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RA GOLDSTEIN I.M., HENNER D., TALHOUK A.;
RL SUBMITTED (MAR-1989) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
```

```
RP SEQUENCE OF 36-193 FROM N.A.
RC TISSUE-BLOOD;
RA RIMSTAD E.;
RL SUBMITTED (JAN-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -!- SUBUNIT: HOMOTRIMER.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL; X14828; G993; -.
DR EMBL; X77317; G452608; -.
DR PIR; S06192; S06192.
DR HSSP; P01375; ITNF.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS00049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT NON_TER 1 1
FT PROPEP <1 37 TUMOR NECROSIS FACTOR.
FT CHAIN 38 193 BY SIMILARITY.
FT DISULFID 106 138 POTENTIAL.
FT CARBOHYD 56 56 R -> S (IN REF. 2).
FT CONFLICT 39 39 E -> A (IN REF. 2).
FT CONFLICT 79 79 T -> N (IN REF. 2).
FT CONFLICT 89 89 H -> Q (IN REF. 2).
FT CONFLICT 115 115 Y -> D (IN REF. 2).
FT CONFLICT 124 124 Q -> L (IN REF. 2).
FT CONFLICT 144 144 Q -> L (IN REF. 2).
FT CONFLICT 174 174 Q -> L (IN REF. 2).
SQ SEQUENCE 193 AA; 21126 MW; AE3F665E CRC32;

Query Match 7.4%; Score 107; DB 1; Length 193;
Best Local Similarity 23.4%; Pred. No. 0.027;
Matches 50; Conservative 36; Mismatches 92; Indels 36; Gaps 10;

QY 84 OGHHAELKPA---GAGAPKAGLEEAPAVTAGLKIFEPAPGEGNSSNSNRKAVQGP 140
DB :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: |||
DB 3 RSHHA-LLPAALRGNRPEE---EQSP---AGPSFNRPLVQTLRSSSQASSNKPVA----- 52
QY 141 TVTQDCLQIADSETPTIQKGSYTFVFWLLSFKRGSALAEKENKILVKETGYFFIYGV 200
DB :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: |||
DB 52 -----HYVANISAP---GOLRWGDSYANALKANGVELKDNQLVPTDGLYISQVL 100
QY 201 Y-----TDKTYAMGHILQKKVHVFGDELSLVTLFR--CIQNNPETLPN---NSCYSAGI 250
DB :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: |||
DB 101 FRGHGCPSTPLFTHTSRIAVS-YQTKVNILSAKSPCHRETPEAKPWTIEPIYQGV 159
QY 251 AKLEEGDELQALAPRENAQISLDGDVTFFGALKL 284
DB :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: ||| :||: |||
DB 160 FOLEKGDRLSAEQINQPEVLDYAESGVVFGIAL 193

RESULT 8
TNFA_PIG
ID TNFA_PIG STANDARD: PRT: 232 AA.
AC P23563;
DT 01-NOV-1991 (REL. 20, CREATED)
DT 01-NOV-1991 (REL. 20, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS SUS SCROFA (PIG).
```

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.

DD 174 SFQREJFEGRHAKFWIEPIIGGVFQJENDKUL-----SAEINLPDILDLFAESGGVI 422
 QY 279 FGALKL 284
 || : |
 Db 227 FGIAL 232

| | |
|------------|--|
| RESULT | 9 |
| TNFA_SHEEP | |
| ID | TNFA_SHEEP STANDARD; PRT; 234 AA. |
| AC | P23383; |
| DT | 01-NOV-1991 (REL. 20, CREATED) |
| DT | 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE) |
| DT | 01-OCT-1994 (REL. 30, LAST ANNOTATION UPDATE) |
| DE | TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN). |
| OS | TNFA. |
| OS | OVIS ARIES (SHEEP). |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; |
| OC | EUTHERIA; ARTIODACTYLA. |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE=LIVER; |
| RX | MEDLINE; 91067496. |
| RA | YOUNG A.J., HAY J.B., CHAN J.Y.C.; |
| RL | NUCLEIC ACIDS RES. 18:6723-6723(1990). |
| RN | [2] |
| RP | SEQUENCE FROM N.A. |
| RC | TISSUE=ALVEOLAR MACROPHAGE; |
| RX | MEDLINE; 92112044. |
| RA | GREEN I.R., SARGAN D.R.; |
| RL | GENE 109:203-210(1991). |
| RN | [3] |
| RP | SEQUENCE FROM N.A. |
| RC | MEDLINE; 92155784. |
| RA | ANDREWS A.E., NASH A.D., BARCHAM G.J., BRANDON M.R.; |
| RL | IMMUNOL. CELL BIOL. 69:273-283(1991). |
| CC | -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN CONDITIONS. |
| CC | -1- SUBUNIT: HOMOTRIMER. |
| CC | -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN EXTRACELLULAR SOLUBLE FORM. |
| CC | -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY PROTEOLYTIC PROCESSING. |
| CC | -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH AND MALNUTRITION. |
| CC | -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY. |
| DR | EMBL; X55966; GI404; - |
| DR | EMBL; X55152; GI406; - |
| DR | EMBL; X56756; G297807; - |
| DR | EMBL; A19163; G512017; - |
| DR | PIR; S13114; S13114. |
| DR | PIR; S20661; S20661. |
| DR | PIR; JH0529; JH0529. |
| DR | HSP; P01375; ITNF. |
| DR | PROSITE; PS00251; TNF_1; 1. |
| DR | PROSITE; PS50049; TNF_2; 1. |
| KW | CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR. |
| FT | PROPEP 1 77 |
| FT | CHAIN 78 234 |
| FT | TRANSMEM 36 56 |
| FT | DISULFID 146 178 |
| FT | CARBOHYD 96 96 |
| FT | CONFLICT 63 63 |
| SQ | SEQUENCE 234 AA; 25536 MW; 8D027AE9 CRC32: |

[illegible]

RX MEDLINE; 93083656.
RA VOIGT C.G., MAUER-FOGY I., ADOLF G.R.;
RL FEBS LETT. 314:85-88(1992).
RN [8]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE; 92129275.
RA ECK M.J., ULTSCH M., RINDERKNECHT E., DE VOS A.M., SPRANG S.R.;
RL J. BIOL. CHEM. 267:2119-2122(1992).
RN [9]
RP VARIANT ALLELE 8.1.
RX MEDLINE; 91139175.
RA ABRAHAM L.J., DU D.C., ZAHEDI K., DAWKINS R.L., WHITEHEAD A.S.;
RL IMMUNOGENETICS 33:50-53(1991).
CC -1- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH
CC IS CYTOTOXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.
CC -1- SUBUNIT: HETEROTRIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
CC OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
CC -1- SUBCELLULAR LOCATION: SECRETED.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL; X01393; G34445; -.
DR EMBL; X02911; G312411; -.
DR EMBL; D00102; D1000508; -.
DR EMBL; M16441; G339740; -.
DR EMBL; D12614; G219912; -.
DR EMBL; M55913; G339743; -.
DR EMBL; Z13026; G37213; -.
DR EMBL; A06316; G412161; -.
DR PIR; A26085; QWHUX.
DR HSSP; P01375; 1TNR.
DR MIN; 153440; -.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS00049; TNF_2; 1.
KW CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL; POLYMORPHISM.
FT SIGNAL 1 34
FT CHAIN 35 205 LYMPHOTOXIN-ALPHA.
FT CARBOHYD 41 41 PARTIAL.
FT CARBOHYD 96 96
FT VARIANT 60 60
FT VARIANT 125 125 T -> N.
FT VARIANT 13 13 T -> P (IN ALLELE 8.1).
FT CONFLICT 13 13 C -> R (IN REF. 6).
SQ SEQUENCE 205 AA; 22297 MW; 42F296D8 CRC32;

Query Match 6.9%; Score 100.5; DB 1; Length 205;
Best Local Similarity 23.6%; Pred. No. 0.1;
Matches 51; Conservative 20; Mismatches 84; Indels 61; Gaps 9;

QY 91 LPAGAGAFKAGLEAPAVTA-----GLKIFPPAPGEGN--SSONSRRKRAVGQPEET 141
DB 29 LPGAQGLPGVGLTSAQTARQHPKMHLSHTLKPAAHLIGDPKQNSLLWRA--NTDRA 86
QY 142 VTQDCLQIADSETPTIOKGSYTFVFWLLSPFKRGSALBEEKNKILVKETGYFFIYGQVLY 201
DB 87 FLQDGFSL-----SNNLLVPTSGIYFYVSQVVF 115
QY 202 TDKIYA----MGHLIQRKVVHVGDELVLVLFRCIQNMPTLPN-----NSCYSAGIA 251
DB 116 SKKAYSPRKATSPPLAHEVQLFSQYFHPVLLSSQKM--VYPLQEPWLHSMYHGAFF 173
QY 252 KLEEGDELQ-----AIPRENAOISLDGDTFFGALKL 284
DB 174 QLTQDQLSTHTDGP-----HLVLSPTVFFGAFAL 205

RESULT 12
TNFA_CAVPO
ID TNFA_CAVPO STANDARD; PRT; 234 AA.
AC P51435;
DT 01-OCT-1996 (REL. 34, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.

OS CAVIA PORCELLUS (GUINEA PIG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HARTLEY; TISSUE=LUNG;
RA YUAN H.T., KELLY F.J., BINGLE C.D.;
RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=DUNKIN-HARTLEY;
RA WHITE A.M., YOSHIMURA T., SMITH A.W., WESTWICK J., WATSON M.L.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION. IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING (BY SIMILARITY).
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL; U39839; G106112; -.
DR EMBL; U77036; G1679724; -.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS00049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 79 BY SIMILARITY.
FT CHAIN 80 234 TUMOR NECROSIS FACTOR.
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DISULFID 147 178 BY SIMILARITY.
SQ SEQUENCE 234 AA; 25793 MW; 00160783 CRC32;

Query Match 6.8%; Score 99; DB 1; Length 234;

Best Local Similarity 22.3%; Pred. No. 0.17;
Matches 55; Conservative 41; Mismatches 85; Indels 66; Gaps 13;

QY 60 CLTVVSFYQVAALQGLASRAELQGHHAELQHPAGAPKAGLEAPAVTAGLKITEPPA 119
DB 32 CLSLFELLVAGATTFLCLLHFGVIGPQEE-QFSSGPPFRPL-----AQTLTLR----- 81
QY 120 PGENSSONSRRKRAVGQPEETVTQDCLQIADSETPTIOKGSYTFVFWLLSPFKRGSAL 179
DB 81 ----SASQNDNDKPVHVAHQAAEELQ-----WL--SKRANALL 115
QY 180 E-----KENKILVKETGYFFIYGQVLYTDK---TY-AMGHLIQRKVVHVGDELVLVTLF 230
DB 116 ANGMGLSDNOLVPSDGLYLIYQVLFKGGCPSYLLTHTVSRLAWS--YPEKVNLLSAI 174
QY 231 R--CIONMPTPLNNSCYS----AGIAKLEEGDELQAIAPRENAOISL-----DGDVT 277
DB 175 KSPCKETPEGAERKPWEPIYLGGVFLQKGRDL-----SAENVLPQYLOFADFSGQI 227
QY 278 FFGALKL 284
DB 228 YFGVIAL 234

RESULT 13
TNFA_HUMAN
ID TNFA_HUMAN STANDARD; PRT; 233 AA.
AC P01375;
DT 21-JUL-1986 (REL. 01, CREATED)
DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)

DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RX SEQUENCE FROM N.A.
RX MEDLINE: 87217060.
RA NEDOSPASOV S.A., SHAKHOV A.N., TURETSKAYA R.L., METT V.A.,
RA AZIZOV M.M., GEORGIEV G.P., KOROKO V.G., DOBRYNIN V.N.,
RA FILIPPOV S.A., BYSTROV N.S., BOLDYREVA E.F., CHUVPILO S.A.,
RA CHUMAKOV A.M., SHINGAROVA L.N., OVCHINNIKOV Y.A.;
RL COLD SPRING HARB. SYMP. QUANT. BIOL. 51:611-624(1986).
RN [2]
RX SEQUENCE FROM N.A.
RX MEDLINE: 85086244.
RA PENNICA D., NEDWIN G.E., HAYFLICK J.S., SEEBURG P.H., DERYNCK R.,
RA PALLADINO M.A., KOHR W.J., AGGARWAL B.B., GOEDDEL D.V.;
RL NATURE 312:724-729(1984).
RN [3]
RX SEQUENCE FROM N.A.
RX MEDLINE: 85137898.
RA SHIRAI T., YAMAGUCHI H., ITO H., TODD C.W., WALLACE R.B.;
RL NATURE 313:803-806(1985).
RN [4]
RX SEQUENCE FROM N.A.
RX MEDLINE: 86016093.
RA NEDWIN G.E., NAYLOR S.L., SAKAGUCHI A.Y., SMITH D.H.,
RA JARRETT-NEDWIN J., PENNICA D., GOEDDEL D.V., GRAY P.W.;
RL NUCLEIC ACIDS RES. 13:6361-6373(1985).
RN [5]
RX SEQUENCE FROM N.A.
RX MEDLINE: 8512190.
RA WANG A.M., CREASEY A.A., LADNER M.B., LIN L.S., STRICKLER J.,
RA VAN ARSDELL J.N., YAMAMOTO R., MARK D.F.;
RL SCIENCE 228:149-154(1985).
RN [6]
RX SEQUENCE FROM N.A.
RX MEDLINE: 86030296.
RA MARVENOUT A., FRANSEN L., TAVERNIER J., DER HEYDEN J., TIZARD R.,
RA KAWASHIMA E., SHAW A., JOHNSON M.J., SEMON D., MUELLER R.,
RA RUYSSCHAERT M.R., VAN VLIET A., FIERIS W.;
RL EUR. J. BIOCHEM. 152:515-522(1985).
RN [7]
RX SEQUENCE FROM N.A.
RX MEDLINE: 93272029.
RA IRIS F.J.M., BOUGUELERET L., PRIEUR S., CATERINA D., PRIMAS G.,
RA PERROT V., JURKA J., RODRIGUEZ-TOME P., CLAVERIE J.-M., DAUSSET J.,
RA COHEN D.;
RL NAT. GENET. 3:137-145(1993).
RN [8]
RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE: 89159409.
RA JONES E.Y., STUART D.I., WALKER N.P.;
RL NATURE 338:225-228(1989).
RN [9]
RX X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).
RX MEDLINE: 91193276.
RA JONES E.Y., STUART D.I., WALKER N.P.;
RL J. CELL SCI. SUPPL. 13:11-18(1990).
RN [10]
RX X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS).
RX MEDLINE: 90008932.
RA ECK M.J., SPRANG S.R.;
RL J. BIOL. CHEM. 264:17595-17605(1989).
RN [11]
RX X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF MUTANT ARG-107.
RX MEDLINE: 98147459.
RA REED C., FU Z.Q., WU J., XUE Y.N., HARRISON R.W., CHEN M.J.,
RA WEBER I.T.;
RL PROTEIN ENG. 10:1101-1107(1997).
RN [12]
RX MUTAGENESIS.

RX MEDLINE: 91184128.
RA OSTADE X.V., TAVERNIER J., PRANGE T., FIERIS W.;
RL EMBO J. 10:827-836(1991).
RN [13]
RX MYRISTOYLATION.
RX MEDLINE: 93018820.
RA STEVENSON F.T., BURSTEN S.L., LOCKSLEY R.M., LOVETT D.H.;
RL J. EXP. MED. 176:1053-1062(1992).
CC -!- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
WITH A WIDE VARIETY OF FUNCTIONS. IT CAN CAUSE CYTOLYSIS OF
CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CONDITIONS.
CC -!- SUBUNIT: HOMOTRIMER.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
EXTRACELLULAR SOLUBLE FORM.
CC -!- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
PROTEOLYTIC PROCESSING.
CC -!- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
AND MALNUTRITION.
CC -!- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: X02910; G37210; -;
DR EMBL: M16441; G339741; -;
DR EMBL: X01394; G37220; -;
DR EMBL: M10988; G339738; -;
DR EMBL: M26331; G339784; -;
DR EMBL: Z15026; G37212; -;
DR PIR: B23784; QWHUN.
DR PIR: A44189; A44189.
DR PIR: S36153; S36153.
DR PDB: 1TNF; 15-JAN-91.
DR PDB: 2TUN; 31-JAN-94.
DR PDB: 1A8M; 17-JUN-98.
DR MIM: 191160; -;
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR;
FT MYRISTYLATION: 3D-STRUCTURE.
FT PROPEP 1 76
FT CHAIN 77 233 TUMOR NECROSIS FACTOR.
FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT LIPID 19 19 MYRISTATE.
FT LIPID 20 20 MYRISTATE.
FT DISULFID 145 177
FT MUTAGEN 105 105 L->S: LOW ACTIVITY.
FT MUTAGEN 108 108 R->W: BIOLOGICALLY INACTIVE.
FT MUTAGEN 112 112 L->F: BIOLOGICALLY INACTIVE.
FT MUTAGEN 160 160 A->V: BIOLOGICALLY INACTIVE.
FT MUTAGEN 162 162 S->F: BIOLOGICALLY INACTIVE.
FT MUTAGEN 167 167 V->A,D: BIOLOGICALLY INACTIVE.
FT MUTAGEN 222 222 E->K: BIOLOGICALLY INACTIVE.
FT CONFLICT 63 63 F -> S (IN REF. 5).
FT TURN 85 86
FT STRAND 89 94
FT TURN 99 100
FT STRAND 104 105
FT TURN 109 110
FT STRAND 112 113
FT TURN 115 116
FT STRAND 118 119
FT STRAND 124 125
FT STRAND 132 138
FT STRAND 140 142
FT STRAND 152 159
FT STRAND 166 174
FT STRAND 190 191
FT STRAND 194 200
FT STRAND 207 212
FT HELIX 215 217
FT STRAND 218 218

This Page Blank (uspto)

Db 176 PCQRETEGAERAKPWYPIYLGQVLEKGRDLSTEI---NLPAYLDFAESGV-YFGII 231
 Qy 283 KL 284
 Db 232 AL 233

RESULT 4
 S24642
 tumor necrosis factor alpha - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 20-Feb-1995 #sequence_revision 20-Feb-1995 #text_change 16-Feb-1997
 C:Accession: I46047; S24642
 R:Clufts, I.; Cleuter, Y.; Kettmann, R.; Burny, A.; Droogmans, L.
 Cytochrome 5, 336-341, 1993
 A:Title: Cloning and characterization of the tandemly arranged bovine lymphotoxin and tu
 A:Reference number: I46046; MUID:94083525
 A:Accession: I46047
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-233 <CL>
 A:Cross-references: EMBL:Z14137; NID:g796; PID:g798
 C:Genetics:
 A:Gene: TNF
 A:Introns: 62/3; 78/1; 94/1
 C:Superfamily: tumor necrosis factor
 C:Keywords: transmembrane protein

Query Match 7.4%; Score 108; DB 2; Length 233;
 Best Local Similarity 22.8%; Pred. No. 0.031;
 Matches 56; Conservative 42; Mismatches 88; Indels 60; Gaps 12;

Qy 58 SC-CLTVVSFYQVAALQGLDASLRAELQGHAEKLPAGAGAPKAGLEAPAVTAGLKIFE 116
 Db 29 SCLCLSLFSLVAGATTFLCLLHFGVIGVQREESPGG-----PSINS----- 72

Qy 117 PPAPGEGNSONSNNKRAVQGPETVTQDCLQLIADSETPTIQKGSYTFVPLLSFKRGS 176
 Db 72 PLVQTLRSSQSSNNKPAV-----HVADINSPGQLRWWDYSANALMA--NGV 117

Qy 177 ALEEKENKILVKETGYFFIYQVLYTDK-----TYAMGHLLQKRVHVFGEDESLVTLFR 231
 Db 118 KLE--DNQLVVPADGLYLIYSQVLFPGQCPSTPLFTHTISRIVS-YQTKVNILSAIK 174

Qy 232 --CIQNPETLP-----NNSCYAGIAKLEEGDELQAIAPRENAQISL-----DGDVTF 278
 Db 175 SPCHRETPAEAKPWYPIYGGVFLQEKGRDL-----SAEINLPDYLDYAESGVY 227

Qy 279 FGALKL 284
 Db 228 FGIAL 233

RESULT 5
 I54490
 tumor necrosis factor - white-footed mouse
 C:Species: Peromyscus leucopus (white-footed mouse)
 C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 23-Feb-1997
 C:Accession: I54490
 R:Crew, M.D.; Filipowsky, M.E.
 Immunogenetics 35, 351-353, 1992
 A:Title: Sequence of the tumor necrosis factor/cachectin (TNF) gene from Peromyscus leu
 A:Reference number: I54490; MUID:92218012
 A:Accession: I54490
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: DNA
 A:Residues: 1-235 <RES>
 A:Cross-references: GB:M92233; NID:g202506; PID:g202507
 C:Genetics:
 A:Gene: P1TNF
 A:Introns: 62/3; 81/1; 97/1

C:Superfamily: tumor necrosis factor

Query Match 7.4%; Score 107.5; DB 2; Length 235;
 Best Local Similarity 22.0%; Pred. No. 0.035;
 Matches 54; Conservative 46; Mismatches 84; Indels 61; Gaps 14;

Qy 60 CLTVVSFYQVAALQGLDASLRAELQGHAEKLPAGAGAPKAGLEAPAVTAGLKIFEPP 118
 Db 32 CLSLFSLVAGATTFLCLLHFGVIGVQREKFP--NLPPIG-----SMAQTILR----- 82

Qy 119 APGENSSONSNNKRAVQGPETVTQDCLQLIADSETPTIQKGSYTFVPLLSFKRGSAL 178
 Db 82 -----SSQNSSDK-----PVAHVANHOVDQLEWLSRGANALL-----ANGM 120

Qy 179 EEKENKILVKETGYFFIYQVLYTDK-----TYA-MGHLLQKRVHVFGEDESLVTLFRCIQ 234
 Db 121 DUKDNLQVLPADGLYLIYSQVLFQGGSSYLLTHTVSRFAVS-YEDKNVLLSAIK--S 177

Qy 235 NNPETLPNNS-----CYSAGIAKLEEGDEL--QLAIPR-----ENNAQISLDGVDVTF 279
 Db 178 PCPKETPEGSELKPNWYPIYLGQVLEKGRDLRAEVNLPKYLDFAESGV-----YF 230

Qy 280 GALKL 284
 Db 231 GVIAL 235

RESULT 6
 SI2606
 tumor necrosis factor alpha precursor - pig
 C:Species: Sus scrofa domestica (domestic pig)
 C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 08-Sep-1997
 C:Accession: SI2606; SI7290; SI8965; I46659
 R:Draws, R.T.; Coffee, B.W.; Prestwood, A.K.; McGraw, R.A.
 Nucleic Acids Res. 18, 5564, 1990
 A:Title: Gene sequence of porcine tumor necrosis factor alpha.
 A:Reference number: SI2606; MUID:91016861
 A:Accession: SI2606
 A:Molecule type: DNA
 A:Residues: 1-232 <DRE>
 A:Cross-references: EMBL:X54001; NID:g2135; PID:g2136
 R:Kuhner, P.; Wuethrich, C.; Peterhans, E.; Pauli, U.
 Gene 102, 171-178, 1991
 A:Title: The porcine tumor necrosis factor-encoding genes: sequence and comparative a
 A:Reference number: SI7289; MUID:91340150
 A:Accession: SI7290
 A:Molecule type: DNA
 A:Residues: 1-232 <KUH>
 A:Cross-references: EMBL:X54859; NID:g2132; PID:g2134
 A:Note: the authors translated the codon GAG for residue 202 as Gly
 R:Choi, C.S.; Mollitor, T.W.; Lin, G.F.; Murtaugh, M.P.
 submitted to the EMBL data library, January 1991
 A:Description: Complete nucleotide sequence of a cDNA encoding porcine tumor necrosis
 A:Reference number: SI8965
 A:Accession: SI8965
 A:Molecule type: mRNA
 A:Residues: 1-232 <CHO>
 A:Cross-references: EMBL:X57321; NID:g2137; PID:g2138
 R:Pauli, U.; Beutler, B.; Peterhans, E.
 Gene 81, 185-191, 1989
 A:Title: Porcine tumor necrosis factor alpha: Cloning with the polymerase chain react
 A:Reference number: I46659; MUID:90034181
 A:Accession: I46659
 A:Status: preliminary; translated from GB/EMBL/DBDJ
 A:Molecule type: mRNA
 A:Residues: 44-232 <PAU>
 A:Cross-references: GB:M29079; NID:g164694; PID:g164695
 C:Genetics:
 A:Introns: 62/3; 78/1; 93/1
 C:Superfamily: tumor necrosis factor
 C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; transmembrane
 F.1-77/Domain: signal sequence #status predicted <SIG>

F:/78-232/Product: tumor necrosis factor alpha #status predicted <MAT>

```

Query Match          7.4%; Score 107; DB 2; Length 232;
Best Local Similarity 22.0%; Pred. No. 0.038;
Matches 54; Conservative 40; Mismatches 86; Indels 66; Gaps 11;

QY 60 CLTVSVFYQVAALQGLASRAELQGHAEKLPAGAPKAGLEAPAVTAGLKIFPPA 119
      |||: || || || || || || || || || || || || || || || || ||
Db 32 CLSFLSFLVAGATLFCLLHFEVIGPQKEFPAGP-----LSI-NPLA 74

QY 120 PEGNSQNSRNRKRAVQGPETVTQDCLQLIADSETPTIQKGSYTFVPWLLSFKR--GS 176
      |||: || || || || || || || || || || || || || || || || ||
Db 75 QGLRSSQTS-----DKPAHVAVNKAEGQ-----LOWQSGYANALLAN 114

QY 177 ALEPEKENKILVKETGYFFIYQVLYTDK-----TYANGHLIQRKKVHVFGDELSLVTFLR 231
      |||: |||: || || || || || || || || || || || || || || || ||
Db 115 GVKLKDNQLVVPDGLYLYSQVLFGRGCPSTNVELTHTISRIAVS-YQTKVNLLSAIK 173

QY 232 --CIQNPETLPNNSCYS-----AGIAKLEGEDELQLAIPRENAQISL-----DGDVTF 278
      |||: || || || || || || || || || || || || || || || || ||
Db 174 SPQRETPGEAGAKPWYPIYLGVFQLEKDDRL-----SAEINLPDYLDFAESQVY 226

QY 279 FGALKL 284
      |||: || || || || || || || || || || || || || || || || ||
Db 227 FGIIAL 232

RESULT 7
S06192
tumor necrosis factor alpha precursor - goat (fragment)
N:Alternate names: cachectin; TNF alpha
C:Species: Capra aegagrus hircus (domestic goat)
C>Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 08-Sep-1997
C:Accession: S06192; S41867
R:Goldstein, I.M.; Henner, D.; Talhouk, A.
submitted to the EMBL Data Library, March 1989
A:Reference number: S06192
A:Accession: S06192
A:Molecule type: mRNA
A:Residues: 1-193 <GOL>
A:Cross-references: EMBL:X14828; NID:g992; PID:g993
R:Rimstad, E.
submitted to the EMBL Data Library, January 1994
A:Reference number: S41867
A:Accession: S41867
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 36-38,'S','Q',40-78,'A','80-88,'N','90-114,'Q','116-123,'D','125-144,'G','145-173,'L'
A:Cross-references: EMBL:X77317; NID:g452607; PID:g452608
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane protein

Query Match          7.4%; Score 107; DB 2; Length 193;
Best Local Similarity 23.4%; Pred. No. 0.03;
Matches 50; Conservative 36; Mismatches 92; Indels 36; Gaps 10;

QY 84 QGHAEKLPA---GAGAPKAGLEAPAVTAGLKIFEPAPGEGNSQNSRNRKRAVQGP 140
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 3 RSHHA-LPAAALNRNRPPEE--EQSP---AGPSFNRLVQTLRSSQASNNKPYA-----52

QY 141 TVTQDCLQLIADSETPTIQKGSYTFVPWLLSFKRGSAALEKENKILVKETGYFFIYQV 200
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 52 -----HVVANISAP-----GQLRWGDSYANALKANGVELKDNQLVVPDGLYLYSQV 100

QY 201 Y-----TDKYAMGHILQIRKKVHVFGDELSLVTFLR--CIQNPETLPN---NSCYAGI 250
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 101 FRHGCGSTPLFTHTSRIAVS-YQTKVNLLSAIKSPCHRETPEAEAKPWYPIYQGV 159

QY 251 AKLEGEDELQLAIPRENAQISLDGDVTFFGALKL 284
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 160 FQLEKGDRLSAINQPEYLDYAESQVYFGIIAL 193

```

```

RESULT 8
S13114
tumor necrosis factor alpha - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 21-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 08-Sep-1997
C:Accession: S13114
R:Young, A.J.; Hay, J.B.; Chan, J.Y.C.
Nucleic Acids Res. 18, 6723, 1990
A:Title: Primary structure of ovine tumor necrosis factor alpha cDNA.
A:Reference number: S13114; MUID:91067496
A:Accession: S13114
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-233 <YOU>
A:Cross-references: EMBL:X55966; NID:g1403; PID:g1404
C:Superfamily: tumor necrosis factor
C:Keywords: transmembrane protein

```

```

Query Match          7.3%; Score 106; DB 2; Length 233;
Best Local Similarity 21.7%; Pred. No. 0.046;
Matches 54; Conservative 41; Mismatches 88; Indels 66; Gaps 11;

QY 58 SC-CLTVSVFYQVAALQGLASRAELQGHAEKLPAGAPKAGLEAPAVTAGLKIFE 116
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 29 SCWCLSLFSFLVAGATTLFCLLHFGVIGPQRESPAGPS-----FN 70

QY 117 PPAPGEGNSQNSRNRKRAVQGPETVTQDCLQLIADSETPTIQKGSYTFVPWLLSFKRGS 176
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 71 RPLVQTLRSSQASNNKPY-----AHVVANISAPGQLRWGDSYANALMA-----N 115

QY 177 ALEPEKENKILVKETGYFFIYQVLY-----TDKYAMGHILQIRKKVHVFGDELSLVTFLR 231
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 116 GVELKDNQLVVPDGLYLYSQVLFGRGCPSTPLTHTISRIAVS-YQTKVNLLSAIK 174

QY 232 --CIQNPETLPN-----NSCYSAGIAKLEGEDELQLAIPRENAQISL-----DGD 275
      |||: |||: || || || || || || || || || || || || || || || || ||
Db 175 SPCHR---ETLEGAERKPWYPIYQGVFQLEKGDRL-----SAEINLPDYLDYAESG 224

QY 276 VTFFGALKL 284
      |||: || || || || || || || || || || || || || || || || ||
Db 225 QVYFGIIAL 233

```

```

RESULT 9
JH0529
tumor necrosis factor alpha precursor - sheep
N:Alternate names: cachectin; TNF alpha
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 08-Sep-1997
C:Accession: JH0529; S48118; S20661
R:Green, I.R.; Sargan, D.R.
Gene 109, 203-210, 1991
A:Title: Sequence of the cDNA encoding ovine tumor necrosis factor-alpha: problems wi
A:Reference number: JH0529; MUID:92112044
A:Accession: JH0529
A:Molecule type: mRNA
A:Residues: 1-234 <GRE>
A:Cross-references: EMBL:X55152; NID:g1405; PID:g1406
A:Experimental source: alveolar macrophage
R:Nash, A.D.; Barcham, G.J.; Brandon, M.R.; Andrews, A.E.
Immunol. Cell Biol. 69, 273-283, 1991
A:Title: Molecular cloning, expression and characterization of ovine TNF-alpha.
A:Reference number: S48118
A:Accession: S48118
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-234 <NAS>
A:Cross-references: EMBL:X56756; NID:g297806; PID:g297807
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane prote

```



```

Db 174 QLTQGDQLSTHTDGP-----HLVLSFSTVFCAAL 205
|||||
RESULT 11
QHUN
tumor necrosis factor alpha precursor - human
N:Alternate names: cachectin; TNF alpha
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1985 #sequence_revision 28-Aug-1985 #text_change 20-Mar-1998
C:Accession: A93585; S36153; A44189; B61478; I53311; S62610; A01646; B23784
R:Newlin, G.E.; Naylor, S.L.; Sakaguchi, A.Y.; Smith, D.; Jarrett-Nedwin, J.; Pennica, D.
Nucleic Acids Res. 13, 6361-6373, 1985
A:Title: Human lymphotoxin and tumor necrosis factor genes: structure, homology and chromosomal location
A:Reference number: A93585; MUID:86016093
A:Accession: A93585
A:Molecule type: DNA
A:Residues: 1-233 <NED>
A:Cross-references: GB:X02910; GB:X02159; NID:g37209; PID:g37210
R:Iris, F.J.M.; Bougueleret, L.; Prieur, S.; Caterina, D.; Primas, G.; Perrot, V.; Jurka
Nature Genet. 3, 137-145, 1993
A:Title: Dense Alu clustering and a potential new member of the NFkappaB family within a
A:Reference number: S36152
A:Accession: S36153
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-233 <IR>
A:Cross-references: EMBL:Z15026; NID:g37211; PID:g37212
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, August 1992
R:Pennica, D.; Nedwin, G.E.; Hayflick, J.S.; Seeburg, P.H.; Derynck, R.; Palladino, M.A.
Nature 312, 724-729, 1984
A:Title: Human tumor necrosis factor: precursor structure, expression and homology to 1
A:Reference number: A93351; MUID:85086244
A:Accession: A93351
A:Molecule type: mRNA
A:Residues: 1-233 <PEN>
A:Note: This protein was isolated from the monocyte-like cell line HL-60 from a promyelocytic leukemia cell line.
R:Wang, A.M.; Creasey, A.A.; Ladner, M.B.; Lin, L.S.; Strickler, J.; Van Arsdel, J.N.;
Science 228, 149-154, 1985
A:Title: Molecular cloning of the complementary DNA for human tumor necrosis factor.
A:Reference number: A44189
A:Accession: A44189
A:Molecule type: mRNA
A:Residues: 1-62, 'S', '64-233 <WAN>
A:Cross-references: GB:M10988; NID:g339737; PID:g339738
R:Fukuda, S.; Ando, S.; Sanou, O.; Tanai, M.; Fujii, M.; Masaki, N.; Nakamura, K.I.; Ando, S.
Lymphokine Res. 7, 173-185, 1988
A:Title: Simultaneous production of natural human tumor necrosis factor-alpha, -beta and -gamma
A:Reference number: A61478
A:Accession: B61478
A:Molecule type: protein
A:Residues: 83-102; 109-119; 121-128, 'X', 130-131; 142-144, 'X', 146, 'XXX', 150-152; 159-174; 180
R:Aggarwal, B.B.; Kohr, W.J.; Hass, P.E.; Moffat, B.; Spencer, S.A.; Henzel, W.J.; Bringman,
J. Biol. Chem. 260, 2345-2354, 1985
A:Title: Human tumor necrosis factor. Production, purification, and characterization.
A:Reference number: A92511; MUID:85130974
A:Contents: annotation; disulfide bond
R:Marnerout, A.; Franssen, L.; Tavernier, J.; Van Der Heyden, J.; Tizard, R.; Kawashima,
Eur. J. Biochem. 152, 515-522, 1985
A:Title: Molecular cloning and expression of human tumor necrosis factor and comparison
A:Reference number: I53311; MUID:86030296
A:Accession: I53311
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-233 <RES>
A:Cross-references: GB:M26331; NID:g339763; PID:g339764
R:Experimental source: U-937 cells
R:Takakura-Yamamoto, R.; Yamamoto, S.; Fukuda, S.; Kurimoto, M.
Eur. J. Biochem. 235, 431-437, 1996
A:Title: O-Glycosylated species of natural human tumor necrosis factor-alpha.
A:Reference number: S62610
A:Accession: S62610
A>Status: preliminary

```

```

A:Molecule type: protein
A:Residues: 77-99 <TAK>
C:Comment: Secreted from mitogen-activated macrophages within 4-24 hours after induction of
detriment to normal cells. It can also act synergistically with interferon gamma
C:Comment: TNF-alpha and -beta (lymphotoxin) are the products of different genes
located on different chromosomes and have different induction kinetics.
C:Genetics:
A:Gene: GDB:TNF; TNFA
A:Cross-references: GDB:120441; OMIM:191160
A:Map position: 6p21.3-6p21.3
A:Introns: 62/3; 78/1; 94/1
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane protein
E:77-233/Product: tumor necrosis factor #status experimental <MAT>
E:145-177/Disulfide bonds: #status experimental

Query Match 6.8%; Score 98.5; DB 1; Length 233;
Best Local Similarity 24.2%; Pred. No. 0.21;
Matches 44; Conservative 34; Mismatches 59; Indels 35; Gaps 9;

Qy 135 VOGPEETVQDCLQI-----ADSETPIQKSYTF-----VPWLLSPKRGSA 178
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 55 VIGPQREFFPRDLSLISPLAQAVRSSRTPSDKPVAVVAVNPQAEGLQWL--NRRANAL 112
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 179 -----EKENKILVKECYFFIYQVLYTDK---TYA-MGHILQKRVHVFGEDEL 228
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 113 LANGVELNDLVVPSEGLYIYQVLFKGGQCPSTHLLTHTISRTAVS-YQTKVNLS 171
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 229 LFR--CIGNMPETLPNNSCYS---AGIAKLEGEDELQALAIAPRENAQISLDGDTFFGAL 282
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 172 AIKSPCQRETEPGAEPWYEPVILGGVQLEKGDRLSAEINRPDYLDFAESGGVYFGII 231
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 283 KL 284
|||
Db 232 AL 233

RESULT 12
A25451
tumor necrosis factor alpha precursor - rabbit
N:Alternate names: cachectin; TNF alpha
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 20-Mar-1998
C:Accession: A25454; A25451; J50727
R:Ito, H.; Shirai, T.; Yamamoto, S.; Akira, M.; Kawahara, S.; Todd, C.W.; Wallace, R.
DNA 5, 157-165, 1986
A:Title: Molecular cloning and expression in Escherichia coli of the cDNA coding for
tumor necrosis factor
A:Reference number: A25454; MUID:86219711
A:Accession: A25454
A:Molecule type: mRNA
A:Residues: 1-234 <ITO>
A:Cross-references: GB:M12845; NID:g165759; PID:g165760
R:Ito, H.; Shirai, T.; Yamamoto, S.; Akira, M.; Kawahara, S.; Todd, C.W.; Wallace, R.
DNA 5, 157-165, 1986
A:Title: Molecular cloning of the gene encoding rabbit tumor necrosis factor.
A:Reference number: A25451; MUID:86219712
A:Accession: A25451
A:Molecule type: DNA
A:Residues: 1-234 <ITO>
A:Note: this sequence differs from that shown in having a Gln inserted between residu
R:Shakhov, A.N.; Kuprash, D.V.; Azizov, M.M.; Jongeneel, C.V.; Nedospasov, S.A.
Gene 95, 215-221, 1990
A:Title: Structural analysis of the rabbit TNF locus, containing the genes encoding T
A:Reference number: JH0309; MUID:91065534
A:Accession: J50727
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-62, 'Q', '63-234 <SHA>
A:Cross-references: GB:M60340; GB:M35326; NID:g165754; PID:g165756
C:Genetics:
A:Introns: 62/3; 80/1; 96/1
C:Superfamily: tumor necrosis factor

```


C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane protein
F:1-81/Domain: signal sequence #status predicted <SIG>
F:82-234/Product: tumor necrosis factor #status predicted <MAT>

Query Match 6.8%; Score 98; DB 2; Length 234;
Best Local Similarity 21.2%; Pred. No. 0.23;
Matches 49; Conservative 46; Mismatches 72; Indels 64; Gaps 13;
QY 112 LKIFPPAPGEGNSSONSNRKRA-----VGGPEETVTQDCL 147
DB 10 VLAEGPLPKKAGPGS--KRCCLSLFSLVAGATLFCLLHPRVIGPQEEESPNNL 67
QY 148 OLIAID-SETPTIOKGSYTF-----VPLLKFKRGSAL-----EKENK 184
DB 68 HLVNVPVQAVMTLSRASRLSDRPLAHVAVNPQVQGLQML--SQANALLANGMKLTDNQ 125
QY 185 ILVKETGYFFIYGOVLTD---KTYA-MGHLIQRKKVHVFGDELSLVTFLR--CIGNMPE 238
DB 126 LVVPADGLYLIYSQVLFSGQGRSYLLTHTVSRFAVS-YPNKVNLLSAIKSPCHRETPE 184
QY 239 TLPNNSCYS---AGIAKLEEGDELQLAIPR-ENAOISLDGDTFFGALKL 284
DB 185 EAPNWPVYPIYGVGFQLEKGRDRLSTEVNQPEYDLAESGVY-YFGIIAL 234

RESULT 13
Fas ligand - human
C:Species: Homo sapiens (man)
C>Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 03-Aug-1998
C:Accession: I38707; JC2340; S57565; 138554
R:Takahashi, T.; Tanaka, M.; Inazawa, J.; Abe, T.; Suda, T.; Nagata, S.
Int. Immunol. 6, 1567-1574, 1994
A:Title: Human Fas ligand: gene structure, chromosomal location and species specificity.
A:Reference number: I38707; MUID:95127360
A:Accession: I38707
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-281 <RES>
A:Cross-references: EMBL:U11821; NID:g595430; PID:g595431
R:Mita, E.; Hayashi, N.; Iio, S.; Takehara, T.; Hijioka, T.; Kasahara, A.; Fusamoto, H.;
Biochem. Biophys. Res. Commun. 204, 468-474, 1994
A:Title: Role of Fas ligand in apoptosis induced by hepatitis C virus infection.
A:Reference number: JC2340
A:Accession: JC2340
A:Molecule type: DNA
A:Residues: 1-281 <MIT>
A:Cross-references: GB:D38122; DBJ:D29820; NID:g601892; PID:d1007898; PID:g1369902
R:Schatzlein, C.E.
submitted to the EMBL Data Library, June 1995
A:Reference number: S57565
A:Accession: S57565
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-281 <SCH>
A:Cross-references: EMBL:X89102; NID:g887455; PID:g887456
R:Allderson, M.R.; Tough, T.W.; Davis-Smith, T.; Braddy, S.; Falk, B.; Schooley, K.A.; Go
J. Exp. Med. 181, 71-77, 1995
A:Title: Fas ligand mediates activation-induced cell death in human T lymphocytes.
A:Reference number: I38554; MUID:95105731
A:Accession: I38554
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-281 <RE2>
A:Cross-references: EMBL:U08137; NID:g624627; PID:g624628
C:Genetics:
A:Gene: FasL
C:Introns: 151/1; 116/3
C:Keywords: glycoprotein; transmembrane protein
F:80-102/Domain: transmembrane #status predicted <TMW>
F:76,184,250,260/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 6.7%; Score 97.5; DB 2; Length 281;
Best Local Similarity 20.7%; Pred. No. 0.32;
Matches 56; Conservative 39; Mismatches 107; Indels 69; Gaps 10;
QY 31 LPRKSPSVR---SSKDGKLLAATLLALLSCCLTVVSFYQVAALQGLDASLR-AELOGH 86
DB 63 LPPLPLPLPKRGNHSTGLCLVMEFMVLVALVGLGLGMFQLFHLQKELAEALRESTSQMH 122
QY 87 HAELKIPAGAGAPKAGLEAPAVTAGLKIFEPAPGEGNSSONSNRKRAVQGPETVTQDC 146
DB 123 TASSLEKQIGHP-----SPP-----PEKELRV 146
QY 147 LOLIADSET---PTIOKGSYTFVPLLSPKRSALKEKENKILVKETGYFFIYGOVLV-- 202
DB 147 AHLTKSNSRSRPLEWEDYIGIV--LL-----SGVYKKGGLVINETGLYFYYSKYVFRG 199
QY 202 -----TDKTYAMGHLIQRKKVHVFGDELSLVTFLFRQNNMPETLPNNSCYSAGIAKL 253
DB 200 QSCNNLPLSHKVMYMRNSKYPQDLVMEGKMSYCTTQ-----MWARSYLGAVFNL 251
QY 254 EGDDELQLAIPRENAOISLDGDTFFGALKL 284
DB 252 TSADHLYNVV-SELSLVNFEESQTFEGLYKL 281

RESULT 14
RGBYD2
translation regulator GCD2 - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein G4615; protein YGR083c
C:Species: Saccharomyces cerevisiae
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 12-Dec-1997
C:Accession: S05809; S64378
R:Padon, C.J.; Hannig, E.M.; Hinnebusch, A.G.
Genetics 122, 551-559, 1989
A:Title: Amino acid sequence similarity between GCN3 and GCD2, positive and negative
A:Reference number: S05809; MUID:89339141
A:Accession: S05809
A:Molecule type: DNA
A:Residues: 1-651 <PAD>
A:Cross-references: EMBL:X15658; NID:g3733; PID:g3734
R:Wedler, H.; Scharfe, M.; Wedler, E.; Wambutt, R.
submitted to the Protein Sequence database, May 1996
A:Reference number: S64356
A:Accession: S64378
A:Molecule type: DNA
A:Residues: 1-651 <WED>
A:Cross-references: EMBL:Z72868; NID:g1323118; PID:e243431; MIPS:YGR083
A:Experimental source: strain S288C
C:Genetics:
A:Gene: SGD;GCD2; GCD12
A:Cross-references: SGD:S0003315; MIPS:YGR083c
A:Map position: 7R
C:Superfamily: translation regulator GCD2
F:573-580/Region: nucleotide-binding motif A (P-loop) #status atypical
F:106/Binding site: phosphate (Ser) (covalent) (by CAMP-dependent kinase) #status pre
F:121/Binding site: phosphate (Thr) (covalent) (by CAMP-dependent kinase) #status pre

Query Match 6.4%; Score 93.5; DB 1; Length 651;
Best Local Similarity 19.1%; Pred. No. 2;
Matches 67; Conservative 55; Mismatches 140; Indels 89; Gaps 15;
QY 2 DDSTEREQSRLTSCLEKREMKKECVSI--LPRKE-----SPSVRSSRSGKLLAA 50
DB 81 EQQKREKQKANKKKQNERVKKSTLFGHLETTERRATILATSVSSPTSITAA 140
QY 51 TLLALLSCCLTVVSFYQVAALQGLDASLR-AELQGHAEKLPAGAGAPKAGLEAPAVTA 110
DB 141 GLMVPVVA-----SALSGSNVLTASSL-----MPVGPNASSTVSASAPASIT 182
QY 111 GLKIFPPAPGEGNSSONSNRKRAVQGPETVTQDCLQLIADSETPTIOKGSYTFYFWL- 170

```
Db 183 TTLPASSAALSAGTSSASTNPTAIQ--QETASSNASDVAKTLASISLEAGEFNVIPIGIS 240
QY 170 -----LSFKRGSALKEE-KE---NK-----ILVKETGYFFIYG----- 198
Db 241 SVIPTVLEQSFQDNSSLSISSKELLNKLHPSILLTSLAHYKIVGSIPTCIAMLEVF 300
QY 198 QVLYTDKTYAMGHLIQRKVVHVFGEDELSL-----VTLFRCIQNNPETL-----PN 242
Db 301 QIVIKDYQTFPGTTLNLTYSILSHQIDLLKARPPLSVTMGNALRWLQKELSLIDPSTPD 360
QY 243 NS-----CYSAG---IAKLEEGDELOLAIPRENAQISLDGDVTFFGALKL 285
Db 361 KAAKKDLCEKIGQFAKEKIELAD--QLIIDNASTQIEESTIIVTYGSSKVL 409
```

```
RESULT 15
S52715
tumor necrosis factor alpha precursor - bovine (fragment)
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-May-1995 #sequence_revision 21-Jul-1995 #text_change 08-Sep-1997
C:Accession: S52715
R:Mertens, B.; Gaigulis, L.
submitted to the EMBL Data Library, March 1995
A:Description: Cloning and sequence analysis of cDNAs encoding bovine CD40 ligand and bo
A:Reference number: S52715
A:Accession: S52715
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-185 <MER>
A:Cross-references: EMBL:248808; NID:g755701; PID:g755702
C:Superfamily: tumor necrosis factor
```

```
Query Match 6.4%; Score 92.5; DB 2; Length 185;
Best Local Similarity 23.0%; Pred. No. 0.52;
Matches 46; Conservative 37; Mismatches 72; Indels 45; Gaps 10;

QY 103 EEAPAVTAGLKIFEPPAPGEGNSSONSRRKRAVQGPETVTQDCLQLIADSETPTOKGS 162
Db 13 EQSP---GGPSINSPLVQTLRSSQASSNKPA-----HVVADINSFGQLRW 57
QY 163 YTFVPWLLSFKRGSALEKENKILVKETGYFFIYGQVLYTDK-----TYAMGHLIQRKV 217
Db 58 DSYANALMA--NGVKLE--DNQLVVPADGLYLYSQVLFRCGCGCPSTPLELTHISRIAV 113
QY 218 HVGDELSLVTLFR--CIQNNPETLP-----NNSCYSAGIAKLEEGDELOLAIPRENAQIS 271
Db 114 S-YQTKVNILSAIKSPCHRETPWAAKWPWEPTYQGGVFQLEKGDRL-----SAEIN 165
QY 272 L-----DGDVTFFGALKL 284
Db 166 LPDYDYAESGGQVIFGIIL 185
```

Search completed: May 13, 1999, 20:59:03
Job time: 2317 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model
Run on: May 13, 1999, 16:35:35 ; Search time 1016.5 seconds
(without alignments)
1926.852 Million cell updates/sec

Title: US-09-212-270-6
Perfect score: 1284
Sequence: 1 accgctcggaattccggg.....aaaaaaaaaaaaaaaa 1284

Scoring table: IDENTITY_NUC

Searched: 2002476 seqs, 762712212 residues

Database : EST.*

- 1: em_est1.*
- 2: em_est2.*
- 3: em_est3.*
- 4: em_est4.*
- 5: em_est5.*
- 6: em_est6.*
- 7: em_est7.*
- 8: em_est8.*
- 9: em_est9.*
- 10: gb_est1.*
- 11: gb_est2.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: gb_est6.*
- 16: gb_est7.*
- 17: gb_est8.*
- 18: gb_est9.*
- 19: gb_est10.*
- 20: gb_est11.*
- 21: gb_est12.*
- 22: gb_est13.*
- 23: gb_est14.*
- 24: gb_est15.*
- 25: gb_est16.*
- 26: gb_est17.*
- 27: gb_est18.*
- 28: gb_est19.*
- 29: gb_est20.*
- 30: gb_est21.*
- 31: gb_est22.*
- 32: em_est10.*
- 33: em_est11.*
- 34: em_est12.*
- 35: em_est13.*
- 36: em_est14.*
- 37: em_est15.*
- 38: em_est16.*
- 39: em_est17.*
- 40: em_est18.*
- 41: em_est19.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB ID | Description |
|------------|-------|-------|--------|-------|--------------------|
| C 1 | 449.2 | 35.0 | 456 | 25 | AA682496 zj18e08.s |
| C 2 | 403.6 | 31.4 | 452 | 24 | AA166695 z08e02.s |

| | | | | | | |
|------|-------|------|-----|----|------------|--------------------|
| C 3 | 376.4 | 29.3 | 378 | 28 | AA906714 | AA906714 ok96d01.s |
| C 4 | 293 | 22.8 | 365 | 10 | R16882 | R16882 yf44c01.s2 |
| C 5 | 275.2 | 21.4 | 377 | 10 | T87299 | T87299 yd89b02.s1 |
| C 6 | 257 | 20.0 | 272 | 20 | HUM307C08B | D79690 Human aorta |
| C 7 | 243.2 | 18.9 | 404 | 30 | AI182472 | AI182472 uc27f12.r |
| C 8 | 210.6 | 16.4 | 458 | 16 | AA254047 | AA254047 val0h03.r |
| C 9 | 207.6 | 16.2 | 383 | 24 | AA422749 | AA422749 vd46a02.s |
| C 10 | 153.6 | 12.0 | 444 | 30 | AI122485 | AI122485 uc62a08.r |
| C 11 | 130.4 | 10.2 | 367 | 10 | R16934 | R16934 yf44c01.r2 |
| C 12 | 52.6 | 4.1 | 434 | 14 | T18718 | T18718 5C04H01-w7 |
| C 13 | 50.8 | 4.0 | 450 | 14 | T18301 | T18301 5C06d10-l7 |
| C 14 | 49 | 3.8 | 588 | 30 | AI118282 | AI118282 ue39f01.x |
| C 15 | 49 | 3.8 | 515 | 31 | AI195176 | AI195176 u16le11.x |
| C 16 | 49 | 3.7 | 728 | 31 | AI256333 | AI256333 u184h05.x |
| C 17 | 48 | 3.7 | 744 | 16 | AA220873 | AA220873 13 Pinus |
| C 18 | 48 | 3.7 | 370 | 16 | HSU46347 | U46347 Human clone |
| C 19 | 48 | 3.7 | 513 | 30 | AI169669 | AI169669 EST215561 |
| C 20 | 47.6 | 3.7 | 231 | 25 | AA693444 | AA693444 ah20b08.s |
| C 21 | 47.6 | 3.7 | 615 | 28 | AA945620 | AA945620 EST201119 |
| C 22 | 47.4 | 3.7 | 596 | 28 | AA945321 | AA945321 EST200820 |
| C 23 | 47.4 | 3.7 | 601 | 28 | AA945586 | AA945586 EST201085 |
| C 24 | 47 | 3.7 | 528 | 30 | AI169809 | AI169809 EST215710 |
| C 25 | 47 | 3.7 | 534 | 30 | AI169826 | AI169826 EST215729 |
| C 26 | 47 | 3.7 | 511 | 30 | AI171464 | AI171464 EST217426 |
| C 27 | 46.6 | 3.6 | 270 | 26 | AA742997 | AA742997 ny20b06.s |
| C 28 | 46.4 | 3.6 | 441 | 28 | AA945161 | AA945161 EST200660 |
| C 29 | 46.2 | 3.6 | 290 | 13 | N81164 | N81164 yw36d03.r1 |
| C 30 | 46 | 3.6 | 464 | 28 | AA945600 | AA945600 EST201099 |
| C 31 | 46 | 3.6 | 425 | 30 | AI169875 | AI169875 EST215784 |
| C 32 | 46 | 3.6 | 543 | 30 | AI169935 | AI169935 EST215845 |
| C 33 | 46 | 3.6 | 427 | 30 | AI169981 | AI169981 EST215897 |
| C 34 | 45.6 | 3.6 | 417 | 23 | AI288116 | AI288116 qv88a03.x |
| C 35 | 45.4 | 3.5 | 398 | 10 | T68838 | T68838 vc30b01.s1 |
| C 36 | 45.4 | 3.5 | 317 | 29 | AA829333 | AA829333 cg07e09.s |
| C 37 | 45.4 | 3.5 | 497 | 30 | AI169775 | AI169775 EST215674 |
| C 38 | 45.2 | 3.5 | 425 | 31 | AI225129 | AI225129 q132h11.x |
| C 39 | 45 | 3.5 | 404 | 28 | AA945000 | AA945000 EST200499 |
| C 40 | 45 | 3.5 | 544 | 28 | C93698 | C93698 Dictyosteli |
| C 41 | 45 | 3.5 | 569 | 30 | AI170010 | AI170010 EST215926 |
| C 42 | 45 | 3.5 | 539 | 30 | AI170036 | AI170036 EST215952 |
| C 43 | 45 | 3.5 | 442 | 31 | AI190896 | AI190896 qd61e08.x |
| C 44 | 44.8 | 3.5 | 260 | 21 | AA279795 | AA279795 zs92a02.s |
| C 45 | 44.8 | 3.5 | 187 | 31 | AI144047 | AI144047 q163c04.x |

ALIGNMENTS

| | | | | | | |
|------------|--|---------------------------|----------|--------------|------|-------------|
| RESULT 1 | AA682496/C | AA682496 | 456 bp | mRNA | EST | 19-DEC-1997 |
| LOCUS | zj18e08.s1 | Soares fetal liver spleen | INFLS S1 | Homo sapiens | cdna | |
| DEFINITION | clone 450662 3', mRNA sequence. | | | | | |
| ACCESSION | AA682496 | | | | | |
| NID | q2669777 | | | | | |
| KEYWORDS | EST | | | | | |
| SOURCE | human. | | | | | |
| ORGANISM | Homo sapiens | | | | | |
| | Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; | | | | | |
| | Primates; Catarrhini; Hominidae; Homo. | | | | | |
| REFERENCE | 1 (bases 1 to 456) | | | | | |
| AUTHORS | Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S., | | | | | |
| | Krizek,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., | | | | | |
| | Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., | | | | | |
| | Theising,B., White,Y., Wyllie,T., Waterston,R. and Wilson,R. | | | | | |
| TITLE | WashU-NCI human EST Project | | | | | |
| JOURNAL | Unpublished (1997) | | | | | |
| COMMENT | | | | | | |

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. Et from Amersham
High quality sequence stop: 456.
Location/Qualifiers

FEATURES
source
1. .456
/organism="Homo sapiens"
/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia) with a modified polylinker; Site_1: Pac I; Site_2: Eco RI; This is a subtracted version of the original Soares fetal liver spleen INFLS library. 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5', AACTGGAAGTAATTAATAGATCTTTTATTTT 3']
double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified p773 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_lib="50662"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 137 a 90 c 85 g 144 t
ORIGIN

Query Match 35.0%; Score 449.2; DB 25; Length 456;
Best Local Similarity 99.3%; Pred. No. 3.5e-101;
Matches 451; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 807 gatettacacattgttccatggtcttcagctttaaagggaagtcgcctagaagaaa 866
|||||
Db 456 GATCTTACACATTGTTCCATGGCTTCTCAGCTTAAAGGGGAAGTGCCCTAGAAGAAA 397
QY 867 aacagataaaatttgtcaagaaactggttacttttttatatatggtcaggttttat 926
|||||
Db 396 AAGAGAATAAATATGCTGCTCAAGAAACTGGTACTTTTATATATGTCAGGTTTTAT 337
QY 927 atactgaagaccctagccatggacatcaatcagaggaaggaagtcacatgtctttg 986
|||||
Db 336 ATACTGATAGACCTACGCCATGGGACATCTAATTCAGAGGAAGAGTCATGCTTTG 277
QY 987 ggaatgaattgactgtgacttggcttgcagtgatgtattcaaaatgctcgaacactac 1046
|||||
Db 276 GGGATGAATTGAGTCTGCTGACTTGTGTCGATGATTCAAAATATGCCCTGAAACACTAC 217
QY 1047 ccaataattcctgctattcagctggcattgcaaaactggaagaggagatgaactccaac 1106
|||||
Db 216 CCATAATTCCTGCTATTCAGCTGGCATTGCAAACTGGAAGAGGAGATGAACCTCAAC 157
QY 1107 ttgcaatccaagaagaaatgcacaaatatacactggatgagatgtcacatttttggtg 1166
|||||
Db 156 TTCAATPACCAAGAGAAATGCACAAATATCAGTGATGATGATGATGATGATGATGATG 97
QY 1167 cattgaactgctgtgacctacttacaccatgtctgtagttattttctccctttctctg 1226
|||||
Db 96 CATTGAACCTGCTGTGACCTACTTACACCATGCTGTGTAGTATTCTCCCTCTCTCTG 37
QY 1227 taccctcaagaagaagaatcactgaataaac 1260
|||||
Db 36 TACTCGAAGAGAAGAATCTAAGCTGAATTAC 3

RESULT 2
AA166695/c
LOCUS
DEFINITION
z08Se02.s1 Stratagene ovarian cancer (#937219) Homo sapiens cDNA
clone 593690 3', mRNA sequence.
AA166695
NID 91745159

KEYWORDS
SOURCE
ORGANISM

EST.
human.
Homo sapiens

Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.

REFERENCE
AUTHORS

1. (bases 1 to 452)
Hallier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M.,
Martin,J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F.,
Theising,B., White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished (1997)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 533 Std Error: 0.00
Seq primer: -40m13 fwd. from Amersham.
Location/Qualifiers
1. .452

FEATURES
source

/organism="Homo sapiens"
/note="Vector: Bluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Cloned unidirectionally. Primer: Oligo dT.
Papillary serous carcinoma, isolated from ascites, 64 year
old caucasian. Average insert size: 0.8 kb; Uni-ZAP XR
vector; -5' adaptor sequence: 5' GAATTCGGCAGAG 3' -3'
adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'
/db_xref="taxon:9606"
/clone_lib="593690"
/sex="female"
/dev_stage="adult, 64 years"
/lab_host="SOLR (kanamycin resistant)"
BASE COUNT 131 a 89 c 86 g 144 t
ORIGIN

Query Match 31.4%; Score 403.6; DB 24; Length 452;
Best Local Similarity 97.8%; Pred. No. 6.4e-90;
Matches 440; Conservative 0; Mismatches 6; Indels 4; Gaps 3;
QY 814 cacattgtccatgcttcagctttaaagggaagtcgcctagaagaaagaa 873
|||||
Db 446 CACATTGTTCCA-GGCTTCTCAGCTTTAAAGGGGAAGTGCCCTGAAG--AAAAAAGAA 390
QY 874 taaaattattggtcaagaactggttacttttttatattatggtcaggttttatatactga 933
|||||
Db 389 TAAATATTGGTC-AAGAACTGGTTACTTTTATATATATGTCAGGTTTATATACTGA 331
QY 934 taagactacagccatggacatcaatcagaggaagaggtccatgtcttgggatga 993
|||||
Db 330 TAAGACCTACGCCATGGGACATCAATTCAGAGGGAAGGTCATGCTTTGGGATGA 271
QY 994 attgactggtgactgttttcgatgtattcaaaatatacctgaaacactaccataa 1053
|||||
Db 270 ATTGAGTCTGCTGACTTTGTTTCGATGATTCAAAATATGCCCTGAACACTACCCATAA 211
QY 1054 ttccgtctattcagctggcattgcaaaactggaagaggagatgaactccacttgcact 1113
|||||
Db 210 TTCCTGCTATTACGTGGCATTGCAAACTGGAAGGAGAGTGAACCTCAACTTGAAT 151
QY 1114 accaagagaaatgcacaaatatacctggtgagatgtcacatttttgggtcattgaa 1173
|||||
Db 150 ACCAAGAGAAAAATGCACAAATATCACTGGATGGAGATGCACATTTTGGTGCATTGAA 91
QY 1174 actgctgacactactacacactgtcttagctatttctccctcttcttcttactct 1233
|||||
Db 90 ACTGCTGTGACCTACTTACACCAATGCTGTAGTATTTTCCCTCCCTTCTCTGTACTCT 31


```

Query Match      20.0%; Score 257; DB 20; Length 272;
Best Local Similarity 97.4%; Pred. No. 9.6e-54;
Matches 265; Conservative 5; Mismatches 1; Indels 1; Gaps 1;

Qy 989 gatgaattgagctggtgactttgtttgagatgtattcaaaatattgcctgaaacataccc 1048
|||||
Db 1 GATGAATTGAGCTGGTGACATTTGTTTCGATGTATTCAAAATATGCTGAAACACTACC 60
|||||

Qy 1049 aataattcctctattcagctggcattgcaaaatggaagagagatgaactcaactt 1108
|||||
Db 61 AATAATTCTCTATTACCTGCGCATTCGCAAACTGGAAGAGGAGATGAATCCCACTT 120
|||||

Qy 1109 gcaataccaagaagaaatgcacaaatatactggatggatgtcacattttttgtgtca 1168
|||||
Db 121 GCAATACCAAGAGAAATGCACAAATATCACTGGATGGAGATGCACATTTTGGTGCA 180
|||||

Qy 1169 ttgaaactgctgtgacctactactaacatgtctgttagtattttccctctctctgta 1228
|||||
Db 181 TTGAAACTGCTGTGACCTACTTACACCATGCTGTGTAGCTATTTTCTCCTCTCTCTGA 240
|||||

Qy 1229 cctctaagaagaagaatctaactgaaataac 1260
|||||
Db 241 CCYTT-NGRAGRAGAAATCTAAGTGAATAATAC 271
|||||

RESULT 7
A1182472 404 bp mRNA EST 08-OCT-1998
LOCUS uc27f12.1 Soares mouse mammary gland NbMMG Mus musculus cDNA clone
DEFINITION 1399247 5', mRNA sequence.
ACCESSION A1182472
NID G3733110
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
AUTHORS Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 404)
Marra,M., Hallier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:910963
Seq primer: -28ml3 rev2 ET from Amersham
High quality sequence stop: 375.
Location/Qualifiers
1. 404
/organism="Mus musculus"
/strain="C57BL/6J"
/Note="Organ: mammary gland; Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5',
TGTTACCAATCTGAAGTGGAGCGCGCGCGGAATGTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldo."
FEATURES
source
/db_xref="taxon:10090"
/clone="1399247"
/clone_lib="Soares mouse mammary gland NbMMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
BASE COUNT 111 a 101 g 98 t
ORIGIN

Query Match      18.9%; Score 243.2; DB 30; Length 404;
Best Local Similarity 78.4%; Pred. No. 2.5e-50;
Matches 304; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

Qy 805 aggatcttacacattgttccattcagcttcaagcttttaaaaggggaagtcgacctagaaga 864
|||||
Db 4 AGGAACCTTACACATTTGTTCCATGGCTTCTCAGCTTTAAAGAGGAAATGCTTGGAGGA 63
|||||

Qy 865 aaagagaataaataatgtcgaagaaactggttacttttttatatatgtcaggtttt 924
|||||
Db 64 GAAGAGAACAAATAGTGGTGAGGCAAAACAGGCTATTCTTCATCTACAGCCAGTTCT 123
|||||

Qy 925 atatactgataagacacctacgccatgggacatcttaattcagagggaagaggtccattgtctt 984
|||||
Db 124 ATACACGGACCCCATCTTTTGTATGGGTGATGTCATCCAGAGGAAGAAAGTACACGTCTT 183
|||||

Qy 985 tggggatgaattgagctggtggaactttgttcgatgtattcaaaatattgcctgaaacact 1044
|||||
Db 184 TGGGACAGGAGTGACCTGTTGACCTCTTCGATGTATTTCAGATATATGCCAAACACT 243
|||||

Qy 1045 accaataattcctgctattcagctggcattgcaaaactgaaagagagatgaactcca 1104
|||||
Db 244 GCCCAACAATCTCTGCTACTCGCTGGCATCGGAGGTGGGAAGAGGATGAGATTCA 303
|||||

Qy 1105 acttgcaataccaagagaaaatgcacaaatatactgagtgagatgtcacattttttgg 1164
|||||
Db 304 GCTTGCAATTCCTCGGGAGATGCACAGATTTTCAGCAACGGAGACGACACTTCTTTGG 363
|||||

Qy 1165 tgcattgaaactgtgtgacctacttac 1192
|||||
Db 364 TGC-CTAAAACTGCTGTAACCTCACTTGC 390
|||||

RESULT 8
AA254047 458 bp mRNA EST 14-MAR-1997
LOCUS val10h03.1 Soares mouse lymph node NbMLN Mus musculus cDNA clone
DEFINITION 722549 5', mRNA sequence.
ACCESSION AA254047
NID g1888612
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
REFERENCE Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
AUTHORS Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
1 (bases 1 to 458)
Marra,M., Hallier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
Unpublished (1996)
Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of MedicineP
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu

```


LOCUS A1122485 444 bp mRNA EST 02-SEP-1998
 DEFINITION uc62a08.r1 Soares mouse mammary gland NDMMG Mus musculus cDNA clone
 1430198 5', mRNA sequence.

ACCESSION A1122485

NID 93522809

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 444)

AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,

Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,

Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,

Theising,B., Wylie,T., Lennon,G., Soares,B., Wilton,R. and

Waterston,R.

TITLE The WashU-HMI Mouse EST Project

JOURNAL Unpublished (1996)

COMMENT

Contact: Marra M/Mouse EST Project

WashU-HMI Mouse EST Project

Washington University School of MedicineP

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: mouseest@watson.wustl.edu

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28ml3 rev2 ET from Amersham

High quality sequence stop: 408.

Location/Qualifiers

1. 444

/organism="Mus musculus"

/strain="C57BL/6J"

/note="Organ: mammary gland; Vector: pT73D-Pac

(Pharmacia) with a modified polylinker; Site 1: Not I;

Site 2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5'

TGTACCAATCTGAAGTGGAGCGGCCGGAATGTTTTTTTTTTTTTTTTTTTT

T 3']; double-stranded cDNA was ligated to Eco RI

adaptors (Pharmacia), digested with Not I and cloned into

the Not I and Eco RI sites of the modified pT73 vector.

RNA provided by Dr. Minoru Ko, Wayne State Univ. Library

constructed and normalized, by Bento Soares and M. Fatima

Bonaldo.

/db_xref="taxon:10090"

/clone="1430198"

/clone_lib="Soares mouse mammary gland NDMMG"

/sex="male"

/tissue_type="mammary gland"

/dev_stage="4 weeks"

/lab_host="DH10B"

BASE COUNT 122 a 116 c 110 g 96 t

ORIGIN

Query Match 12.0%; Score 153.6; DB 30; Length 444;
 Best Local Similarity 72.8%; Pred. No. 3.4e-28;
 Matches 198; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

QY 996 tgagctggtgactgttttcgatgtattcaaaaatagctgaacacatacctaataatt 1055

DB 1 TGAGCTGTGTGACCTGTTCCTCCAGATGATTTCAGAAATATGCCAAACACATGCCCAACAATT 60

QY 1056 ctgctattcagctggcattgcaaaactggaagagagatgaactcccaacttgaatac 1115

DB 61 CTGCTACTCGCTGGCATCGGAGGCTGGAGAGAGAGATGAGATTCAGCTTGCAATTC 120

QY 1116 caagagaaatgcacaaatatactctggatggagatgtcacattttttgtgcatgaac 1175

DB 121 CTCGGGAGAAATGCACAGATTCACGCGAGGAGACACACCTTCCTTGTGCTTAAAC 180

QY 1176 tgctgtgacctacttaccacatgtctgtagctatttttctccctctctctgtacctctaa 1235

DB 181 TGCTGTAACTACTTGCCTGGAGTGGGTATCCCTTCCTCTCTCTGTGTACCTCGA 240

QY 1236 gaagaaagaatctaactgaataacaaaaaa 1267

DB 241 GGGAGAAACAGACGACTGGAATAAATAAAGA 272

RESULT 11

LOCUS R16934

DEFINITION Yf44c01.r2 Homo sapiens cDNA clone 129696 5'.

ACCESSION R16934

NID 9770544

KEYWORDS EST.

SOURCE human clone-129696 library-Soares fetal liver spleen LNFUS

vector-pT73D (Pharmacia) with a modified polylinker host=DH10B

(ampicillin resistant) primer-M13RP1 Rsite1-Pac I Rsite2-Eco RI

Liver and spleen from a 20 week-post conception male fetus. 1st

strand cDNA was primed with a Pac I - oligo(dT) primer [5'

ACTGGAAGAAATTAATAAGATCTTTTTTTTTTTTTTTTTTTTT 3'], double-stranded

cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac

I and cloned into the Pac I and Eco RI sites of the modified pT73

vector. Library went through one round of normalization. Library

constructed by Bento Soares and M. Fatima Bonaldo.

ORGANISM Homo sapiens

Eucaryotae; Metazoa; Chordata; Vertebrata; Gnathostomata; Mammalia;

Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 367)

AUTHORS

Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,

Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,

Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,

Trevasakis,E., Waterston,R., Williamson,A., Wohlmann,P. and

Wilson,R.

TITLE The WashU-Merck EST Project

JOURNAL Unpublished (1995)

COMMENT

Contact: Wilson RK

WashU-Merck EST Project

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@watson.wustl.edu

High quality sequence stops: 292

Source: IMAGE Consortium, LNL

This clone is available royalty-free through LNL; contact the

IMAGE Consortium (info@image.llnl.gov) for further information.

FEATURES

Location/Qualifiers

1..367

/organism="Homo sapiens"

/clone="129696"

BASE COUNT 112 a 83 c 99 g 69 t 4 others

ORIGIN

Query Match 10.2%; Score 130.4; DB 10; Length 367;
 Best Local Similarity 93.6%; Pred. No. 1.8e-22;
 Matches 190; Conservative 0; Mismatches 7; Indels 6; Gaps 5;

QY 525 accaggtgcccgcctcgaaggcgacctggccagctccggcagagctgcggggccacc 584

DB 110 ACAAGTGTGGCCCTCGCAAGGGNACCTGGCCAGCTCCGGCAGAGCTGCAGGGCCACC 169

QY 585 acgcggaagaagctccacagagcagagcccccaagccgacct-ggagaaactcca 643

DB 170 ACGGGAGGAAGCTGCCAGCAGGAGCAGAGCCCCCAAGCCCGCTGGGAGGAAGCTCCA 229

QY 644 gctg-tcacccgggact--gaaatctttgaaccaccagctccca-ggagaaggcaactc 699

DB 230 GCTGTTCACCGGGGACTTGAAATCTTTGAACACCAGCTCCAGGGAGAGGCAATTC 289

Search completed: May 13, 1999, 19:44:05
Job time: 11310 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 01:59:41 ; Search time 2254.51 Seconds
(without alignments)
2037.503 Million cell updates/sec

Title: US-09-212-270-6

Perfect score: 1284

Sequence: 1 accggtccggaattccggg.....aaaaaaaaaaaaaaaaaaaa 1284

Scoring table: IDENTITY_NUC

Searched: 808301 seqs, 1788773984 residues

Database :

GenEmbl.*

1: gb_bal.*

2: gb_ba2.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl1.*

9: gb_pl2.*

10: gb_pri.*

11: gb_pr2.*

12: gb_pr3.*

13: gb_ro.*

14: gb_st.*

15: gb_sy.*

16: gb_un.*

17: gb_vi.*

18: gb_htg.*

19: em_ba.*

20: em_fun.*

21: em_hum1.*

22: em_hum2.*

23: em_in.*

24: em_om.*

25: em_or.*

26: em_ov.*

27: em_pat.*

28: em_ph.*

29: em_pl.*

30: em_ro.*

31: em_sy.*

32: em_un.*

33: em_vl.*

34: em_htg.*

35: em_sts.*

36: gb_bal.*

37: gb_ba2.*

38: gb_pl1.*

39: gb_pl2.*

40: gb_pri.*

41: gb_pr2.*

42: gb_pr3.*

43: gb_sts.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|-------------|
|------------|-------|-------------|--------|----|-------------|

ALIGNMENTS

RESULT 1

G30081/c

LOCUS

DEFINITION

ACCESSION

NID

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

JOURNAL

COMMENT

04-OCT-1996

STS

DNA

365 bp

sequence tagged site.

human STS SHGC-36171, sequence tagged site.

human.

human.

human.

human.

human.

human.

Primer A: GTACAGAGAAAGGAGGAAATAGC
Primer B: ATGCACAAATATCACTGGATGG
STS size: 105
PCR Profile:

Initial incubation: 94 degrees C for 90 seconds

Denaturation: 94 degrees C for 15 seconds
Annealing: 62 degrees C for 23 seconds
Polymerization: 72 degrees C for 30 seconds
PCR Cycles: 30

Thermal Cycler: Perkin Elmer 9600

Protocol:
Template: 25 ng
each 1 uM
Primer:
dNTPs: each 200 uM
Taq Polymerase: 0.05 units/ul
Total Vol: 10 ul

Buffer: MgCl2: 2.5 mM
KCl: 50 mM
Tris-HCl: 20 mM
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from R16802
-- Washington University/Merck EST sequence.

FEATURES

source
1. .365
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="13"
43. .147

primer_bind 43. .67

primer_bind 94 a 70 c 116 t 9 others
BASE COUNT
ORIGIN

Query Match 22.8%; Score 293; DB 43; Length 365;
Best Local Similarity 94.7%; Pred. No. 1.1e-62;
Matches 341; Conservative 0; Mismatches 14; Indels 5; Gaps 4;

Qy 917 caggtttatata-ctgataagacctag-cctggagacatc--taattcagagagaag 972
Db 360 CAGGTTTATANACTGATAGAGCCCGCCATGGGACATCCTAANTCCAGAGAAGAA 301
Qy 973 ggtt-ccatgtcttggggatgaattgagctgtggtgactttgttcgatgtattcaaaata 1031
Db 300 GGNCCCATGCTTTGGGGATGAANTGAGTCTGGTGACCTTTGTTTCGATGTATTCAAAATA 241
Qy 1032 tgcctgaacactaccaataattcctgctattcagctggcattgcaaaactggaagaag 1091
Db 240 TGCCCTGAANACTACCCAATAANTCCCTGCTANTCAGCTGGCATTGCAAAACTGGAAGAAG 181
Qy 1092 gagatgaactccaacttgcaatacaagagaaaatgcacaaatatcactgagatgagatg 1151
Db 180 GAGATGAATCCACTTGCAATNCAGAGAGAAATGACAAATATCACTGGATGGAGATG 121
Qy 1152 tcacatttttgggtgactgaaactgctgacacctactacacattcctgtagctattt 1211
Db 120 TCACATTTTGGTGCATTGAAACTGCTGTGACCTACTTANACCATGCTCTAGCTATT 61
Qy 1212 tcctccctttctgtacctctaagaagaagaatctactgaataatacaaaaaa 1271
Db 60 TCCTCCCTTCTCTGTACCTTAAGAGAGAAAGATCTAACTGAAATACCAAAAAA 1

RESULT 2

I66494/c I66494 7218 bp DNA PAT 23-DEC-1997
LOCUS Sequence 14 from patent US 5670367.
DEFINITION I66494
ACCESSION I66494
NID g2724471

KEYWORDS
SOURCE
ORGANISM

Unknown.
Unclassified.

REFERENCE 1 (bases 1 to 7218)
Dorner, F., Scheiflinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES Location/Qualifiers
source 1. .7218
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 4.3%; Score 55.6; DB 6; Length 7218;
Best Local Similarity 2.7%; Pred. NO. 0.0011;
Matches 10; Conservative 221; Mismatches 145; Indels 0; Gaps 0;

Qy 48 gagaagactttgaaattcttcaaaaaactgaaagtgaatgaggaagacattgagcaa 107
Db 1426 RRR 1367
Qy 108 tccaatcgagggttaaatgccagcaaacctactgtacagttaggggttagagatgcagagaag 167
Db 1366 RRR 1307
Qy 168 gcagaaagagagaaattcaggataaactctctcaggggtgagccacccctgccatgag 227
Db 1306 RRR 1247
Qy 228 tgcacgcagacatcaacaaacagagataacaggaatgatccattccctgtgtcactt 287
Db 1246 RRR 1187
Qy 288 attctaaagcccaactcacaagtccaagttagatgtagatgtagatgcacagagaagg 347
Db 1186 RRR 1127
Qy 348 agcagtccacgccttactcttgccttaagaaagagagaataaactgaagagagtgtg 407
Db 1126 RRR 1067
Qy 408 ttccatcctccacac 423
Db 1066 ATCGCAAGCTCCCTCG 1051

RESULT 3

TSU88240 625 bp mRNA INV 04-MAR-1997
LOCUS Trichinella spiralis hypothetical ORF 2.20 mRNA, partial cds.
DEFINITION U88240
ACCESSION
NID g1857757

KEYWORDS
SOURCE
ORGANISM

Trichinella spiralis.
Trichinella spiralis
Eukaryotae; mitochondrial eukaryotes; Metazoa; Nematoda;
Adenophorea; Enoplia; Trichocephalida; Trichocephalatinia;
Trichinelloidea; Trichinellidae; Trichinella.

REFERENCE 1 (bases 1 to 625)
Polvere, R.I. and Despommer, D.D.
AUTHORS Submitted (04-FEB-1997) Environmental Science, Columbia University,
JOURNAL 630 W. 168th Street, New York, NY 10032, USA
FEATURES Location/Qualifiers
source 1. .625
/organism="Trichinella spiralis"
/db_xref="taxon:6334"
/clone="2.20"
CDS <1. .507
/note="hypothetical ORF 2.20"
/codon_start=1

| BASE COUNT | 650 a | 415 c | 450 g | 563 t |
|------------|-------|-------|-------|-------|
| ORIGIN | | | | |

[illegible]

| | Query Match | 3.4% | Score 43.4; | DB 8; | Length 1148; | |
|----|--------------------------|--|-----------------|--------|--------------|------|
| | Best Local Similarity | 62.4%; | Pred. No. 0.99; | | | |
| | Matches 68; Conservative | 0; | Mismatches 41; | Indels | 0; | Gaps |
| | | | | | | 0; |
| QY | 1176 | tgcgtgacacttaccacatgtcgtgtagctattttcccttccttctcttacctcta | 1235 | | | |
| | | | | | | |
| Db | 1010 | TGCTGTAGTGTAAATGGCAGTGTTACTATTATTAAGCGCCCTTTAAGAACCTCCA | 1069 | | | |
| | | | | | | |
| QY | 1236 | gaagaagaatctcaactgaatacacaaaaaaaaaaaaaaaaaaaaa | 1284 | | | |
| | | | | | | |
| Db | 1070 | AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA | 1118 | | | |

| Result No. | Query | | | Description | | |
|------------|-------|-------|--------|-------------|--------|---------------------|
| | Score | Match | Length | DB | ID | |
| 1 | 244.5 | 16.9 | 250 | 2 | 075888 | 075888 homo sapien |
| 2 | 107 | 7.4 | 233 | 4 | 018779 | 018779 bos taurus |
| 3 | 107 | 7.4 | 232 | 10 | 035853 | 035853 mus musculus |
| 4 | 106.5 | 7.3 | 234 | 4 | Q28320 | Q28320 capra hircu |
| 5 | 102.5 | 7.1 | 391 | 2 | 075910 | 075910 homo sapien |
| 6 | 93.5 | 6.4 | 157 | 2 | 043647 | 043647 homo sapien |
| 7 | 92 | 6.3 | 156 | 10 | 062336 | 062336 mus musculus |
| 8 | 90.5 | 6.2 | 301 | 9 | P74823 | P74823 sphingomona |
| 9 | 87.5 | 6.0 | 233 | 4 | 077510 | 077510 papio hamad |
| 10 | 87.5 | 6.0 | 927 | 9 | 087015 | 087015 pseudomonas |
| 11 | 87.5 | 6.0 | 377 | 10 | 054693 | 054693 mus musculus |
| 12 | 87 | 6.0 | 216 | 10 | 070332 | 070332 mesocricetu |
| 13 | 86.5 | 6.0 | 1229 | 8 | 049749 | 049749 arabidopsis |
| 14 | 86.5 | 6.0 | 233 | 10 | 035734 | 035734 maxmota mon |
| 15 | 85.5 | 5.9 | 231 | 3 | P90611 | P90611 toxoplasma |
| 16 | 85.5 | 5.9 | 233 | 4 | 077764 | 077764 macropus eu |
| 17 | 85.5 | 5.9 | 846 | 9 | P71133 | P71133 chlamydia p |
| 18 | 84 | 5.8 | 344 | 2 | 060513 | 060513 homo sapien |
| 19 | 84 | 5.8 | 461 | 3 | Q21419 | Q21419 caenorhabdi |
| 20 | 83 | 5.7 | 426 | 3 | 045371 | 045371 caenorhabdi |
| 21 | 83 | 5.7 | 656 | 10 | 088461 | 088461 rattus norv |
| 22 | 82.5 | 5.7 | 760 | 1 | Q99126 | Q99126 ustilago ma |
| 23 | 82.5 | 5.7 | 900 | 1 | 074925 | 074925 schizosacch |
| 24 | 82.5 | 5.7 | 693 | 3 | 061920 | 061920 caenorhabdi |
| 25 | 82.5 | 5.7 | 371 | 9 | 033322 | 033322 mycobacteri |
| 26 | 82 | 5.7 | 1015 | 3 | Q09938 | Q09938 caenorhabdi |
| 27 | 82 | 5.7 | 838 | 4 | 019056 | 019056 papio hamad |
| 28 | 82 | 5.7 | 1286 | 8 | 023351 | 023351 arabidopsis |
| 29 | 81.5 | 5.6 | 562 | 9 | P73196 | P73196 synechocyst |

```

Db   32 CLSLFELLVAGATTLCLLNFVGIVGQDEKFPNG-----LPLISSMAQTU--- : 79
QY   119 APGEENSSQRNRKAQGVPEETVTQCLOLIADSETPTIQGSYTFVPWLLSFKRGSAL 178
Db   79 -----SSQNSSDKPVAAHVANHQVEQLWSQRSNALLANG-----M 117
QY   179 EEKENKLIVKETGYFFIYGVOVLTYDK----TYAMGHILIOKKVHHVFGEDELSTLTFR--C 232
Db   118 DLKDNLVFPADGLLYYSOVLVKFGQGCPDYVLTTHTVSRAIS-YOEKVNLUSAVKSPC 176
QY   233 IQNNPTELP-----NNSCYSGIAKLEEGDEL--QLAIPR-----ENAQSISLDGDVTFPGA 281
Db   177 PKDTPEGAELKPWEPIYLGCVFLKEKGDLQAESVNLPKYLDFAESGOV-----YFGV 229
QY   282 LKL 284
Db   230 IAL 232

RESULT      4
Q28320          PRELIMINARY;           PRT;    234 AA.
ID AC Q28320;
DT 01-NOV-1996 (TREMBREL. 01, CREATED)
DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE TNF-ALPHA.
OS CAPRA HIRCUS (GOAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUETHERIA;
CN ARTIODACTYLTA; RUMINANTIA; PECORA; BOVIDAE; BOVIDAE; CAPRINAE; CAPRA.
[1]
RP SEQUENCE FROM N.A.
RA TAKAKURA H., MORI Y., TATSUMI M.;
RL SUBMITTED (JUL-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; D86587; G1483165; -
DR PROSITE; PS00251; TNF_1; 1.
DR PFAM; PF00229; TNF; 1.
SQ SEQUENCE 234 AA; 25519 MW; C6424744 CRC32;

Query Match              7.3%; Score 106.5; DB 4; Length 234;
Best Local Similarity    22.2%; Pred. No. 0.15;
Matches 53; Conservative 37; Mismatches 104; Indels 45; Gaps

QY   58 SC-CLTVSYQVAALOGDLASLAERAELOGHHAEKLPAGACAPKAGLEEPAVNTAGLKIFE 116
Db   29 SCWCLSLFSELLVAGATTLCLLHFVGIGPORE-----EQSP---AGPSFNK 72
QY   117 PPAPGEGNSSONSRNKRAVGPETVTQCLOLIADSEPTTIQGSYTFVPWLLSFKRGS 176
Db   73 PLVOTLRSSQASSNPVA-----HVVANISAP-----GOLRGSDSYANAALKAN 116
QY   177 ALEEKENKIIVKETGYFFIYGVOVLY-----TDKTVMAGHLIOKKVHHVFGEDELSTLTFR 231
Db   117 GVLEKDNQLVPTDGLYLIYSQVLFRHGCHPCSTPLETLTISRAYS-VQTKVNILSAIK 175
QY   232 --CIQNNPETLP-----NNSCYSGIAKLEEGDELQAIAPRNAQISLDGDVTFFGAIKL 284
Db   176 SPCHRETPEGAEAKPWEPYIYQGVFLKEKGRDSLAEINOPELVDAESGGVVFEGIAL 234

RESULT      5
Q75910          PRELIMINARY;           PRT;    391 AA.
ID AC Q75910;
DT 01-NOV-1998 (TREMBREL. 08, CREATED)
DT 01-NOV-1998 (TREMBREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE ECTODYSPLASIN-A ADA ISOFORM II.
GN ED1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUETHERIA.
```

| | | | | | |
|----|--------------------------|--|----------------|--------|--------------------|
| | Query Match | 6.3%; | Score 92; | DB 10; | Length 156; |
| | Best Local Similarity | 19.5%; | Pred. No. 1.4; | | |
| | Matches 32; Conservative | 41; | Mismatches | 57; | Indels 34; Gaps 7; |
| QY | 124 | NSSONSINKRAVGPEETVTQDCIQLIADSETPIIQKSYTFVPWLLSFKRGSALKEEN | 183 | | |
| | | : : : : : : | : : : : : : | | |
| Db | 3 | SSONSQDKPVAVHVANHQVEEQLEWISQRANALLANG-----MDUKDN | 46 | | |
| | | : : : : : : | : : : : : : | | |
| QY | 184 | KILYKETGYFTFYGOVLYTDR---TYAMGHLIOKKVHVFGEDELSVLTLFR--CIQNMP | 237 | | |
| | | : : : : : : | : : : : : : | | |
| Db | 47 | QLVVPADGLYLIVYSOVLFKGCGCPDYVLLTHTVSRFAIS-YQEKNVLLSAVKSPCKDTP | 105 | | |
| | | : : : : : : | : : : : : : | | |
| QY | 238 | ETLP-----NNSCYSGAGIAKEEGDEL--QLAI PR-----ENAQI | 270 | | |
| | | : : : : : : | : : : : : : | | |
| Db | 106 | EGAECLKPWIEPYLGGYGFOLEKGPOLGAENVNLPKYLDFAESGVQ | 149 | | |
| | | : : : : : : | : : : : : : | | |

QY QY
184 KILVKETGEFFIYGVNLYTDK----TYAMGHLILQRKKVHVGDEUSLVTLFR--CLQNMP 433
::: |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 47 QLVPVADGLYLVSQVLFGSGCGCPDVVLLTHTVTSRFAIS-YQEKNVLLSAVKSPCKDTP 105

QY 238 ETLP-----NNSCYSAGIAKLEEGDEL--QLAIPR-----ENAQI 270

Db 106 EGAEIKPWYEPIYLGGVFOLEKGDSLSAEVNLPKYLDLDFAESGV 149
::: |::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RP SEQUENCE FROM N.A.
RC STRAIN=588;
RA MEDLINE: 96196177.
RX YAMAZAKI M., THORNE L., MIKOLAJCZAK M., ARMENTROUT R.W., POLLOCK T.J.;
RT "linkage of genes essential for synthesis of a polysaccharide capsule
RT in *Sphingomonas* strain 588.";
RJ J. BACTERIOL. 178:2676-2687(1996).
DR EMBL: U51197; G1314571; -.
KW HYPOTHETICAL PROTEIN.
SQ SEQUENCE 301 AA; 31458 MW; E1C4B275 CRC32;

Query Match 6.2%; Score 90.5; DB 9; Length 301;
Best Local Similarity 23.8%; Pred. No. 4.3;
Matches 51; Conservative 24; Mismatches 74; Indels 65; Gaps
OV 33 RKSPSPVRSKDGKGLAATLLALLSCCLTTVVSF---YQVAALQGDLASLRAELOGHHAE 89

```

Db 114 RGETAVTLAEKLAGLLRAGGYAKPIVSVEVSVNYITVLGQVITAGLQPVDRGVHVS 173
QY 90 KLPAGACAPKAGLEAPAVTAGLKIFEPAPGCGNSQNSRNRKRAVQGPETVTQDCLQL 149
Db 174 EIIARAGGLURADAADFVLT-----ADGTSKLNKQLAQGGPE----- 214
QY 150 IADSETPTIQKGYTFVFWLLSFKRSALAEKENKILVKETGYFFIYGQV-----LVT 202
Db 214 -----QPVVTPGDKLFP-----EVEFYIYGQVNAQGVYAIRT 248
QY 203 DKT-----YAMGHLL-----QRKKVHVGDEL 226
Db 249 DMTLRRALAQGGGLTPAGSSKRVKVSROGQEIKL 282

RESULT 9
O77510
ID O77510 PRELIMINARY; PRT; 233 AA.
AC O77510;
DT 01-NOV-1998 (TREMBREL. 08, CREATED)
DT 01-NOV-1998 (TREMBREL. 08, LAST SEQUENCE UPDATE)
DE TUMOR NECROSIS FACTOR ALPHA PRECURSOR.
OS PAPIO HAMADRYAS URSINUS (CHACMA BABOON).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;
OC PRIMATES; CATARRHINI; CERCOPITHECIDAE; CERCOPITHECINAE; PAPIO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 98147379.
RA HAUDEK S.B., REDL H., SCHLAG G., GIROIR B.P.;
RT "Complementary DNA (cDNA) sequence of baboon tumor necrosis factor
alpha."
RL MOL. IMMUNOL. 34:1041-1042(1997).
DR EMBL; AF019963; G3417555; -.
DR PROSITE; PS00251; TNF_1; 1.
KW SIGNAL.
FT CHAIN 1 76 POTENTIAL.
FT TAG 77 233 TUMOR NECROSIS FACTOR ALPHA.
SQ SEQUENCE 233 AA; 25658 MW; E38CF6E8 CRC32;

```

```

Query Match 6.0%; Score 87.5; DB 4; Length 233;
Best Local Similarity 19.9%; Pred. No. 5;
Matches 49; Conservative 38; Mismatches 92; Indels 67; Gaps 9;

QY 61 LTVSVFYQVAALOGDLSLRAELQGHAEKLPAGACAPKAGLEAPAVTAGLKIFEPAP 120
Db 33 LSLFSLVAGATTFLCFLHFGVIGPQREFFPKDP-----SLISPLAQ 75
QY 121 GEGNSSQNSRNRKRAVQGPETVTQDCLQIADSETPTIQKGYTFVFWLLSFKRSAL-- 179
Db 76 AVRSSRTSDKPVVHVANPQAEGLQ-----WL--NRRANALLA 114
QY 179 ---EERENKILVKETGYFFIYGQVLTDTKYAMGHLIQRKKVH---VFGDELSLVTFLR 231
Db 115 NGVELTDNLVLPSEGLIYQVLFKGGCPSNVLHTTISRIVASYQTKNVLSAIAK 174
QY 232 --CIQNPETPLNNSCYS-----AGIAKLEEGDELQAIAPRENAQISL-----DGDVTF 278
Db 175 SPQRETPGEAEAKPWYEPIYLGGVFQLEKGRDL-----SAEINLPDYLDFAESGQVY 227
QY 279 FGALKL 284
Db 228 FGIAL 233

RESULT 10
O87015
ID O87015 PRELIMINARY; PRT; 927 AA.
AC O87015;
DT 01-NOV-1998 (TREMBREL. 08, CREATED)
DT 01-NOV-1998 (TREMBREL. 08, LAST SEQUENCE UPDATE)

```

```

DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE FIMV.
GN PSEUDOMONAS AERUGINOSA.
OS BACTERIA; PROTEOBACTERIA; GAMMA SUBDIVISION; PSEUDOMONAS GROUP;
OC PSEUDOMONAS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PAOLI;
RA SEMMLER A., WHITCHURCH C.B., MATTICK J.S.;
RT "Pseudomonas aeruginosa twitching motility gene fimV."
RL SUBMITTED (MAR-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; U93274; G3237312; -.
SQ SEQUENCE 927 AA; 98172 MW; 2642488E CRC32;

Query Match 6.0%; Score 87.5; DB 9; Length 927;
Best Local Similarity 24.1%; Pred. No. 32;
Matches 51; Conservative 24; Mismatches 80; Indels 57; Gaps 8;

QY 3 DSTEREQSLTSLCKKREEM--KLKECVSILPRKESPSVRSKDGKLAATLLALLSCC 60
Db 327 DSTRRENEELQSRMDLQSLDKLQIQL-----KDA----- 360
QY 61 LTVSVFYQVAALOGDL-----ASLRAELQGHAEKLPAGACAPKAGLE----- 104
Db 360 -----QLAKLOGQLGAEQGAAPNAAALPDASQPNAAQAAPQPGTPAAAAPTAPAG 412
QY 104 EAPAVTAGLKIFEPAP-CEGNSQNSRNRKRAVQGPETVTQDCLQIADSETPTIQKS 162
Db 413 EAPAPAPQPPVAPPAPAEKPPAPVAPVAPVQAAEQAPSPFLDELLANPLWLAVIGS 472
QY 163 --YTFVFWLLSFKRSALAEKEN-KILVKETG 191
Db 473 ALLALIVLLMILSRNAQKEEAQAFAADTG 504

RESULT 11
O54693
ID O54693 PRELIMINARY; PRT; 377 AA.
AC O54693;
DT 01-JUN-1998 (TREMBREL. 06, CREATED)
DT 01-JUN-1998 (TREMBREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)
DE EDA PROTEIN HOMOLOG.
GN TABBY.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97449184.
RA FERGUSON B.M., BROCKDORFF N., FORMSTONE E., NGUYEN T., KRONMILLER J.E.,
RA ZONANA J.;
RT "Cloning of tabby, the murine homolog of the human EDA gene: evidence
for a membrane-associated protein with a short collagenous domain."
RL HUM. MOL. GENET. 6:1589-1594(1997).
RN [2]
RP SEQUENCE OF 1-132 FROM N.A.
RC STRAIN-129;
RA BROCKDORFF N., ZONANA J., FORMSTONE E., FERGUSON B.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; AF004435; G2660700; -.
DR EMBL; AF004434; G2660698; -.
SQ SEQUENCE 377 AA; 40002 MW; AEA7ED8 CRC32;

```

```

Query Match 6.0%; Score 87.5; DB 10; Length 377;
Best Local Similarity 18.0%; Pred. No. 10;
Matches 39; Conservative 40; Mismatches 79; Indels 59; Gaps 9;

QY 92 PAGAGAPKA--GLEEAPAVTAGLKIFEPAPGCGNSQNSRNRKRAVQGP-----EETVTQD 145
Db 01-NOV-1998 (TREMBREL. 08, LAST SEQUENCE UPDATE)

```



```
Query Match      6.0%; Score 86.5; DB 10; Length 233;
Best Local Similarity 21.2%; Pred. No. 6.6;
Matches 55; Conservative 39; Mismatches 73; Indels 93; Gaps 14;

QY 60 CLTVSYQVAALQGLDASLAEALQGHAEK-----LPAGAGAPKAGLEAPAVTAGLKIF 115
   |||: || || | : | : || |
Db 32 CLSLFSLLVAGATFTCLLHFGVIGPQREFFLNFLPSQAQMLTLR----- 80
   |||: || || | : | : || |

QY 116 EPPAPGEGNSNSNRKRAVGPEETV---TODCLQLIADSETPTIOKGSYTFVPWLLSF 172
   |||: || || | : | : || |
Db 80 -----SSQNNDK-----PVAHVAKNDEQLV-----WL--S 107
   |||: || || | : | : || |

QY 173 KRGSAL-----EKENKILVKETGYFFIYQVLYTDK---TYA-MGHLIORKKVVHFGDE 223
   |||: || || | : | : || |
Db 108 RRANALLANGMELIDNQLVVPANGLYLVISQVLFKGQGPCSYVLLTHTVSRFAVS-YQDK 166
   |||: || || | : | : || |

QY 224 LSLVTLFRCTONMPETLPNNS-----CYSAGIAKLEGGDELQLAIPRENAQIS 271
   |||: || || | : | : || |
Db 167 VNLSAIAK-----SPCPKESLEGAEPKWPYEPYILGGVFELQKGDRL-----SAENV 213
   |||: || || | : | : || |

QY 272 L-----DGDVTFEGALKL 284
   |||: || || | : | : || |
Db 214 LPSYLDFAESGQVTFGVIAL 233
   |||: || || | : | : || |

RESULT 15
P90611 PRELIMINARY; PRT; 231 AA.
AC P90611;
DT 01-MAY-1997 (TREMBLREL. 03, CREATED)
DT 01-MAY-1997 (TREMBLREL. 03, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE H4 GENE.
GN H4.
OS TOXOPLASMA GONDII.
OC EUKARYOTA; ALVEOLATA; APICOMPLEXA; COCCIDIA; EIMERIIDA; SARCOCYSTITDAE;
OC TOXOPLASMA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BK;
RA NOCKEMANN S.;
RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL; Y09782; E284082; -.
SQ SEQUENCE 231 AA; 25984 MW; CC32AB49 CRC32;

Query Match      5.9%; Score 85.5; DB 3; Length 231;
Best Local Similarity 26.7%; Pred. No. 7.9;
Matches 50; Conservative 21; Mismatches 79; Indels 37; Gaps 8;

QY 18 KREEMKLEKCVSILPRKESPSVRSKDGKLLAATLLALLSCCLTVSYFYQVAALQGLA 77
   |||: || || | : | : |||
Db 43 KREPFLYSRVTML-----RPTVRS-----LJSLGLTVILY---LALTGSAD 82
   |||: || || | : | : |||

QY 78 SLRAELQCHAE-----KLPAAGAGAPKAGLEAPAVTAGLKIFPPAPGEGNSNSNRKR 133
   |||: || || | : | : |||
Db 83 ALASHLSRHRMEAGRRYMDTNDVESAGRQSEPMEDADRAEHGCAQTQ--SEMKEFOEE 140
   |||: || || | : | : |||

QY 134 AVQGPETVTODCLQIADSETP-----TIQKGSYTFVPWLLSFKRGSALAEKENKI-LVK 188
   |||: || || | : | : |||
Db 141 IKGVEETKHED-----DPEMTRLMTVEKQESKFNKMAKSQSFSRIEELGGSISFLT 194
   |||: || || | : | : |||

QY 189 ETGYFFI 195
   |||: || || | : | : |||
Db 195 ETGVMTI 201
   |||: || || | : | : |||
```

Search completed: May 14, 1999, 21:39:56
Job time: 7990 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 10:14:51 ; Search time 129.53 Seconds
(without alignments)
1864.900 Million cell updates/sec

Title: US-09-212-270-6
Perfect score: 1284
Sequence: 1 accgctcggaattcccg...aaaaaaaaaaaaaaaa 1284

Scoring table: IDENTITY_NUC

Searched: 240622 seqs, 94065609 residues

Database : N_Geneseq_34.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------|--------------------|
| 1 | 1098.4 | 85.5 | 1100 | 1 V30934 | Homo sapiens neutr |
| 2 | 1095.2 | 85.3 | 1101 | 1 V39985 | Human T cell surfa |
| 3 | 1088.8 | 84.8 | 1093 | 1 V58894 | cDNA encoding tumo |
| 4 | 1071.2 | 83.4 | 1109 | 1 V39984 | Human T cell surfa |
| 5 | 886 | 69.0 | 987 | 1 V39986 | Human T cell surfa |
| 6 | 302.8 | 23.6 | 338 | 1 V30935 | Homo sapiens neutr |
| 7 | 275.6 | 21.5 | 617 | 1 V39987 | Mouse T cell surfa |
| 8 | 259.4 | 20.2 | 509 | 1 V30936 | Homo sapiens neutr |
| 9 | 189.4 | 14.8 | 497 | 1 V30937 | Homo sapiens neutr |
| 10 | 16.8 | 3.6 | 1328 | 1 V58754 | Human secreted pro |
| 11 | 44.6 | 3.5 | 2705 | 1 V05164 | Human growth arrs |
| 12 | 44.4 | 3.5 | 2639 | 1 V52934 | Pig transforming g |
| 13 | 42 | 3.3 | 2239 | 1 V01527 | Wheat soluble star |
| 14 | 41.6 | 3.2 | 1638 | 1 T06480 | Cystathionine gamm |
| 15 | 41.4 | 3.2 | 2676 | 1 Q02819 | cDNA sequence enco |
| 16 | 41.4 | 3.2 | 2671 | 1 Q03303 | Entire porcine tra |
| 17 | 41.4 | 3.2 | 1013 | 1 Q12431 | Fragment D of urat |
| 18 | 41.4 | 3.2 | 2669 | 1 Q56925 | Pig TGF-beta-3. Nu |
| 19 | 41 | 3.2 | 525 | 1 V41453 | Nucleotide sequenc |
| 20 | 41 | 3.2 | 2323 | 1 V59524 | Human secreted pro |
| 21 | 41 | 3.2 | 1378 | 1 V59706 | Human secreted pro |
| 22 | 40.6 | 3.2 | 6020 | 1 Q06648 | Plasminogen gene f |
| 23 | 40.6 | 3.2 | 6010 | 1 Q11998 | Human plasminogen |
| 24 | 40.6 | 3.2 | 971 | 1 Q73760 | RP-8 programmed ce |
| 25 | 40.6 | 3.2 | 971 | 1 T35704 | Coding sequence fo |
| 26 | 40.4 | 3.1 | 2064 | 1 N92583 | Sequence of human |
| 27 | 40.4 | 3.1 | 2064 | 1 T66701 | Human serum albumi |
| 28 | 40.4 | 3.1 | 3214 | 1 T62360 | Human origin of re |
| 29 | 40.4 | 3.1 | 3214 | 1 T73287 | Human origin of re |
| 30 | 40.4 | 3.1 | 2205 | 1 Q04680 | Human presenilin i |
| 31 | 40.4 | 3.1 | 1373 | 1 V52969 | Human G1 protein i |
| 32 | 40.2 | 3.1 | 819 | 1 V63198 | cDNA from clone ga |
| 33 | 39.8 | 3.1 | 2001 | 1 T59700 | PTH-like peptide D |
| 34 | 39.8 | 3.1 | 1454 | 1 T39913 | Plant CAD enzyme D |
| 35 | 39.8 | 3.1 | 1474 | 1 V33914 | Plant CAD enzyme D |
| 36 | 39.8 | 3.1 | 1508 | 1 V07665 | Maize Bx1 cDNA (DI |
| 37 | 39.6 | 3.1 | 940 | 1 V07281 | Arabidopsis STZ po |
| 38 | 39.6 | 3.1 | 1627 | 1 V40486 | Human secreted pro |
| 39 | 39.4 | 3.1 | 2916 | 1 Q11655 | TSW S RNA. DNA co |
| 40 | 39.4 | 3.1 | 2215 | 1 Q70882 | Tomato spotted wil |
| 41 | 39.4 | 3.1 | 5058 | 1 V35114 | Mouse WRN helicase |
| 42 | 39.4 | 3.1 | 921 | 1 V40521 | Homo sapiens CH27- |
| 43 | 39.4 | 3.1 | 3899 | 1 V63192 | cDNA from clone dt |

ALIGNMENTS

```
RESULT 1
V30934
ID V30934 standard; DNA; 1100 BP.
AC V30934;
DT 11-SEP-1998 (first entry)
DE Homo sapiens neutrokin alpha protein gene.
KW neutrokin alpha; cell proliferation; differentiation; migration;
KW cytotoxicity; cell death; treatment; tumour; infection; inflammation;
KW wound healing; immunodeficiency; autoimmune disease; graft rejection;
KW fibrotic disorder; haematopoiesis; sepsis; shock; malaria; HIV; AIDS;
KW acquired immune deficiency syndrome; rheumatoid arthritis; silicosis;
KW cachexia; detection; diagnosis; drug screening; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT CDS 147..1004
/*tag= a
/product= neutrokin alpha
W09818921-A1.
07-MAY-1998.
25-OCT-1996; U17957.
25-OCT-1996; WO-U17957.
(HUMA-) HUMAN GENOME SCI INC.
Ebner R, Ni J, Yu G;
WPI: 98-272216/24.
P-PSDB: W58391.
New isolated human Neutrokin alpha - used to develop products for
diagnosis and treatment of e.g. tumours, infections,
immunodeficiencies or autoimmune diseases
Claim 2; Fig 1; 104pp; English.
The sequence is that encoding neutrokin alpha.
Neutrokin alpha (NA) polypeptides modulate cell proliferation,
differentiation, migration, cytotoxicity and cell death.
They can be used to treat e.g. tumour and tumour metastasis, infections
by bacteria, viruses and other parasites, immunodeficiencies, graft
inflammatory diseases, lymphadenopathy, autoimmune diseases, graft
versus host disease and to stimulate peripheral tolerance, destroy some
transformed cell lines, mediate cell activation and proliferation, and
are functionally linked as primary mediators of immune regulation and
inflammatory responses. Such activity is useful for immune enhancement
or suppression, myeloprotection, stem cell mobilisation, acute and
chronic inflammatory control and treatment of leukaemia. They can also
be used to stimulate wound healing and to treat fibrotic disorders
including liver cirrhosis, osteoarthritis and pulmonary fibrosis. They
can also be used to regulate haematopoiesis, by regulating the activation
and differentiation of various haematopoietic progenitor cells, e.g. to
release mature leukocytes from the bone marrow following chemotherapy,
and in stem cell mobilisation. NA may also be used to treat sepsis. NA
antagonists can be used to prevent septic shock, inflammation, cerebral
malaria, activation of the HIV virus, graft-host rejection, bone
resorption, rheumatoid arthritis and cachexia (wasting or malnutrition).
They can also be used to treat e.g. autoimmune diseases such as multiple
sclerosis and insulin-dependent diabetes and inflammatory and infectious
diseases such as silicosis, and sarcoidosis, idiopathic pulmonary
fibrosis, idiopathic hyper-eosinophilic syndrome, endotoxic shock,
atherosclerosis, histamine-mediated allergic reactions and immunological
disorders including late phase allergic reactions, chronic urticaria, and
atopic dermatitis by inhibiting chemokine-induced mast cell and basophil
degranulation and release of histamine. IGE-mediated allergic reactions
such as allergic asthma, rhinitis and eczema, inflammatory pulmonary
diseases, rheumatoid arthritis, inflammation, degenerative and
inflammatory arthropathies, aplastic anaemia, myelodysplastic syndrome,
subepithelial basement membrane fibrosis or adult respiratory distress
syndrome. The products can also be used for detection, diagnosis and
drug screening.
Sequence 1100 BP; 343 A; 259 C; 250 G; 248 T;
```

44 39.4 3.1 2025 1 V59809 Human secreted pro
45 39.4 3.1 2346 1 V59693 Human secreted pro


```
Qy 341 gaaagggagcagtcacgccttactcttctgcttaagaaagagaagaatgaaactgaag 400
Db 181 GAAAGGGAGCAGTCACGCCTTACTCTTTCCTTAAGAAAGAGAGAAATGAATGAGG 240
Qy 401 gagtggtttccatctctccacggaagaaagccctctgtccgatctccaaagacgga 460
Db 241 GAGTGTGTTTCCATCTCTCCACGGAAGAAAGCCCTCTGTCCGATCTCTCCAAAGACGGA 300
Qy 461 aagctgtcgtgctcaactctgctgacgactgtgtcttctgtctgctcactcagtggtgtct 520
Db 301 AAGCTGTGCTGTGCAACTCTGCTGTGCACTGTCTGTCTGTGCTCACGGTGGTGCT 360
Qy 521 ttctaccaggtggcgccctgcaaggggacctggccagcctccgggagagctgcagggc 580
Db 361 TTCACAGGTGGCGCCCTGCAAGGGACCTGGCCACGCTCCGGGCAGAGCTGCAGGGC 420
Qy 581 caccacgcggaagagctgcagcaggaagcagagagcccccagggccgctggaggaagct 640
Db 421 CACCACGGGAGAGCTGCCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCT 480
Qy 641 ccagctgtcacgcggaactgaaactcttgaaccacagctccagagagaaggaactcc 700
Db 481 CCAGCTGTACCGGGGACTGAAATCTTTGAACCAACAGCTCCAGGAGAGGCAACTCC 540
Qy 701 agtcagacagcagaaataagctgcccgttcagggtccagagaaacagtcactcaagac 760
Db 541 AGTCAGAACACAGAAATAAGCGTGCCTTCAGGGTCCAGAGAAACAGTCACCTCAAGAC 600
Qy 761 tgcctgaactgattgcagacagtgaaacacaaactatcacaaagagatcttaacattt 820
Db 601 TGCCTGCAACTGATTGTCAGACAGTGAACACACCACTATACAAAAAGGATCTTACACATT 660
Qy 821 gtccatggtcttcagctttaaagggggaagtgccctagagaaagagaataaata 880
Db 661 GTTCATGGCTTCTCAGCTTTAAAGGGGAGTGCCCTAGAGAGAAAGAGAAATAAATA 720
Qy 881 ttggtcaagaaactgggtactttttatatatggtcaggttttatatactgataagacc 940
Db 721 TTGGTCAAGAAACTGGTTACTTTTTATATATGTCAGGTTTATATACGTATGATGAGCC 780
Qy 941 tacgccatggacatctaatcagaggaagagtgccatgtctgttgggagtaattgagt 1000
Db 781 TACGCCATGGGACATCTAATTCAGAGGAAGAGGTCCATGTCTTTGGGGATGAATTGAGT 840
Qy 1001 ctggtgactgtttctgattatcaaatatgctgaaacactacccaataattctctgc 1060
Db 841 CTGGTGACTTTGTTTCGATGTATTCAAAATATGCGTGAACACACTACCCAAATATCTCTGC 900
Qy 1061 tattcagctggcattgcaaaactggaagagagatgaactcccaactgcaatccaaga 1120
Db 901 TATTCAGCTGGCATGTGCAAACTGGAGAGAGAGATGAATCTCCAATCTGCAATCAAGA 960
Qy 1121 gaaatgcacaaatatactggatggagatgcatcatttttttgggtgcatgaaactgctg 1180
Db 961 GAAATGCACAAATATCACTGGATGGAGATGTCACATTTTGGTGATTTGAAACTGGCTG 1020
Qy 1181 tgacctactcacacatgctgtgagctatttctcctctctctctctctctctctctctctct 1240
Db 1021 TGACCTACTTACACATGCTGTAGCTATTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1080
Qy 1241 aagaatcctaactgaaataac 1260
Db 1081 AAGAATCTAATGAAATAC 1100
```

RESULT 3

```
V58894
ID V58894 standard; cDNA; 1093 BP.
AC V58894;
DT 07-JAN-1999 (first entry)
DE cDNA encoding tumour necrosis factor homologue TL5 protein.
KW Tumour necrosis factor homologue TL5; vaccine; chronic;
KW acute inflammation; arthritis; septicemia; autoimmune disease;
```

```
KW inflammatory bowel disease; psoriasis; transplant rejection;
KW graft vs. host disease; infection; stroke; ischaemia;
KW acute respirator disease syndrome; restenosis; brain injury; AIDS;
KW bone disease; cancer; lymphoproliferative disorder; atherosclerosis;
OS Alzheimer's disease; ss.
FH Key Location/Qualifiers
CDS 154..1011
FT /*tag= a
FT /product= TL5
PN EP-869180-A1.
PD 07-OCT-1998.
PF 01-APR-1998; 302526.
PR 03-DEC-1997; US-984396.
PR 02-APR-1997; US-041797.
PA (SWIK ) SMITHKLINE BEECHAM CORP.
PI Hurlie MR, Young PR;
DR WPI; 98-508494/44.
DR P-PSDB; W73043.
PT New tumour necrosis factor homologue, TL5 - useful for diagnosis and
PT treatment of Alzheimer's disease, AIDS and cancer
PS Claim 2; Page 17; 23pp; English.
CC The present sequence encodes a tumour necrosis factor homologue TL5
CC polypeptide sequence. TL5 polypeptides and antibodies are useful for
CC identifying compounds which agonise and antagonise TL5, and these can be
CC administered for treatment to inhibit TL5 activity (antagonist) or
CC enhance TL5 activity (agonist). Gene therapy using the expression system
CC can also be used to enhance TL5 activity. Diseases or susceptibility to a
CC disease can be diagnosed by determining the presence or absence of a
CC mutation in the TL5 protein. TL5 polynucleotides are useful for locating
CC genes associated with disease by hybridisation to chromosomes. TL5
CC polypeptides and polynucleotides can be used, especially to raise an
CC immune response (i.e. as vaccines) for the treatment of chronic and acute
CC inflammation, arthritis, septicemia, autoimmune diseases (e.g.
CC inflammatory bowel disease, psoriasis), transplant rejection,
CC graft vs. host disease, infection, stroke, ischaemia, acute respiratory
CC disease syndrome, restenosis, brain injury, AIDS, bone diseases, cancer
CC (e.g. lymphoproliferative disorders), atherosclerosis, and Alzheimers
CC disease.
SQ Sequence 1093 BP; 332 A; 261 C; 252 G; 248 T;
```

```
Query Match 84.8%; Score 1088.8; DB 1; Length 1093;
Best Local Similarity 99.8%; Pred. No. 1.8e-270;
Matches 1090; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
Qy 174 agggaaaaattcaggataaactctctgaggggtgagccaagccctgcatgtagtcacg 233
Db 2 ACGAGAAATTCAGATAAATCTCTCTGAGGGGTGAGCCCAAGCCCTGCCATGTAGTGCAGC 61
Qy 234 caggacatcaacaacacagataacaggaaatgatccattccctgtgtgctatttcta 293
Db 62 CAGGACATCAACAAACACAGATAACAGGAATGATCCATTCCCTGTGTCTACTATTCTA 121
Qy 294 aagggcccaactcacaagtccaagtagtagatgtagtgcactccacagaagaggagcagt 353
Db 122 AAGGCCCAACTCTTCAAGTTCAAGTAGTATGGATGACTCCACAGAAAGGGAGCAGT 181
Qy 354 cagccttactcttgccttaagaaaagagagaataaactgaagagtgtgtttcca 413
Db 182 CACGCTTACTTCTTGCTTTAAGAAAGAGAGAAATGAAACTGAGGAGTGTGTTTCCA 241
Qy 414 tcctcccaaggaagaaagccctctgtccgatactcccaagagcggaaagtgtggtg 473
Db 242 TCCTCCCAACGGAAGAAAGCCCTCTGTCCGATCCTCCAAAGACGGAAGAGTGTGGCTG 301
Qy 474 caactgtctgtggcactgtgtctgtcctcactcagcaggtgtgtcttcttaccagg 533
Db 302 CAACCTTGCTGCTGGCACTGTGTCTTCTGCTCCTCACGGTGTGTCTTTTCTACCAAG 361
Qy 534 cgcctctgaaggggacctgccagcctccgggagagctcagagtcagggccaccacgggaga 593
Db 362 CGGCCTGCAAGGGGACCTGCCAGCCTTCCGGGAGAGCTCGAGGCGCCACCACCGCGGAGA 421
```


Query Match 21.5%; Score 275.6; DB 1; Length 617;
 Best Local Similarity 74.5%; Pred. No. 8.7e-62;
 Matches 347; Conservative 0; Mismatches 119; Indels 0; Gaps 0;

Qy 802 aaaaagagataaaatattggtcaagaaagactgttacttttttatatatgttcagggt 921
 Db 74 AAACAGAACTTACACATTTGTTCCATGCTTCTCAGCTTAAAGAGGAAATGCCTTGA 133

Qy 862 agaaaagagataaaatattggtcaagaaagactgttacttttttatatatgttcagggt 921
 Db 134 GGAGAAAGAGAAACAAATAGTGGTGGAGCAACAGGCTATTCTTCATCTACAGCCAGGT 193

Qy 922 ttatatactgataagacctacgcatgggacatactaatccagagagaaaggtccatgt 981
 Db 194 TCTATACAGGACCCCACTTTGCTATGGTCATGTCATCCAGAGAGAAAGTACAGT 253

Qy 982 ctttgggatgaattgagctggtgactttgttggatgtattcacaatatgcctgaac 1041
 Db 254 CTTTGGGACGAGCTGAGCTGGTGGACCTGTTCCGATGTATTTCAGATATGCCCAAAAC 313

Qy 1042 actaccataattcctctctattcagctggcattgcaaaactggaagagagatgaact 1101
 Db 314 ACTGCCCAACAAATTCCTGCTACTCGCTGGCATCGGAGCTGGAGAGGAGATGAGAT 373

Qy 1102 ccaacttgcaatccaagagaaaaatgcacaaatactggatggatgtcacattttt 1161
 Db 374 TCAGCTTCAATTCCTCGGGAGAAATGCAGATTTTCACGACGAGACACCTTCTT 433

Qy 1162 tgggtgattgaactgctgtagcctactacaccatgtctgtagtattttcctccctt 1221
 Db 434 TGGTCCCTAAACATGCTGTAACCTACTGCTGAGTGGAGTCCCTTCCCTCGTCTT 493

Qy 1222 cctgtactctaaagaaagaatcctaactgaaatacaaaaaa 1267
 Db 494 CTCTGACTCTCGAGGAGAAACAGACGACTGGAAAACTAAAGA 539

RESULT 8
 V30936 standard; DNA; 509 BP.
 ID V30936;
 AC V30936;
 DE Homo sapiens neurotrophin alpha gene related clone HSLAH84R.
 DT neurotrophin alpha; cell proliferation; differentiation; migration;
 KW cytotoxicity; cell death; treatment; tumour; infection; inflammation;
 KW wound healing; immunodeficiency; autoimmune disease; graft rejection;
 KW fibrotic disorder; haematopoiesis; sepsis; shock; malaria; HIV; AIDS;
 KW acquired immune deficiency syndrome; rheumatoid arthritis; silicosis;
 KW cachexia; detection; diagnosis; drug screening; HSLAH84R; ss.
 OS Homo sapiens.
 PN W09818921-A1.
 PD 07-MAY-1998.
 PF 25-OCT-1996; U17957.
 PR 25-OCT-1996; WO-017957.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ebner R, Ni J, Yu G;
 DR WPI; 98-272216/24.
 PT New isolated human Neurotrophin alpha - used to develop products for
 diagnosis and treatment of e.g. tumours, infections,
 PT immunodeficiencies or autoimmune diseases
 PS Claim 21; Page 86; 104pp; English.
 CC The sequence is that of the neurotrophin alpha related clone HSLAH84R.
 CC Neurotrophin alpha (NA) polypeptides modulate cell proliferation,
 CC differentiation, migration, cytotoxicity and cell death.
 CC They can be used to treat e.g. tumour and tumour metastasis, infections
 CC by bacteria, viruses and other parasites, immunodeficiencies,
 CC inflammatory diseases, lymphadenopathy, autoimmune diseases, graft
 CC versus host disease and to stimulate peripheral tolerance, destroy some
 CC transformed cell lines, mediate cell activation and proliferation, and
 CC are functionally linked as primary mediators of immune regulation and
 CC inflammatory responses. Such activity is useful for immune enhancement
 CC or suppression, myeloprotection, stem cell mobilisation, acute and

CC chronic inflammatory control and treatment of leukaemia. They can also
 CC be used to stimulate wound healing and to treat fibrotic disorders
 CC including liver cirrhosis, osteoarthritis and pulmonary fibrosis. They
 CC can also be used to regulate haematopoiesis, by regulating the activation
 CC and differentiation of various haematopoietic progenitor cells, e.g. to
 CC release mature leukocytes from the bone marrow following chemotherapy,
 CC and in stem cell mobilisation. NA may also be used to treat sepsis, NA
 CC antagonists can be used to prevent septic shock, inflammation, cerebral
 CC malaria, activation of the HIV virus, graft-host rejection, bone
 CC resorption, rheumatoid arthritis and cachexia (wasting or malnutrition).
 CC They can also be used to treat e.g. autoimmune diseases such as multiple
 CC sclerosis and insulin-dependent diabetes and inflammatory and infectious
 CC diseases such as silicosis, and sarcoidosis, idiopathic pulmonary
 CC fibrosis, idiopathic hyper-eosinophilic syndrome, endotoxic shock,
 CC atherosclerosis, histamine-mediated allergic reactions and immunological
 CC disorders including late phase allergic reactions, chronic urticaria, and
 CC atopic dermatitis by inhibiting chemokine-induced mast cell and basophil
 CC degranulation and release of histamine. IGE-mediated allergic reactions
 CC such as allergic asthma, rhinitis and eczema, inflammatory pulmonary
 CC diseases, rheumatoid arthritis, inflammation, degenerative and
 CC inflammatory arthropathies, aplastic anaemia, myelodysplastic syndrome,
 CC subepithelial basement membrane fibrosis or adult respiratory distress
 CC syndrome. The products can also be used for detection, diagnosis and
 CC drug screening.
 SQ Sequence 509 BP; 130 A; 76 C; 115 G; 142 T;

Query Match 20.2%; Score 259.4; DB 1; Length 509;
 Best Local Similarity 82.3%; Pred. No. 1.2e-57;
 Matches 339; Conservative 0; Mismatches 64; Indels 9; Gaps 4;

Qy 881 ttggtcaagaactggtacttttttatatatgctcaggttttatatactgataagacc 940
 Db 4 TCGGCANAGNAACTGGTTACTTTTATATATGCTCAGGTTTATATCTAGTAAAGACC 63

Qy 941 tacgcatggagacatctaatccagaggaaggtccatgtctttggggatgaattgagt 1000
 Db 64 TAGGCGCATGGGACATCTAGTTTCAGAGGAAGAGGTCCATGCTTTGGGGATGAATTGAGT 123

Qy 1001 ctggtgactttggttcgatgtattcaaaaatgctgcaaacactaccataaattccctgc 1060
 Db 124 CTGCTGACTTTGTTTCGATGATTTCAAAATATGCTGAAACACTACCCCAATTAATTCCTGC 183

Qy 1061 tattcagctggcattgcaaaactgaa-gaaggagatgaactccaactgcaatccaacaa 1119
 Db 184 TATTGAGCTGGCATTCGAAACTGAGAGGAGATGAATCCCACTGCAATACCAAGG 243

Qy 1120 agaaaatgcacaaatattcact-ggatggagatgt---cacattttttggtgctgaac 1175
 Db 244 GGAAATGCACAAATATTACATGGGATGAGATGTTACATTTTGGGTGCAATGAAAC 303

Qy 1176 tgcgtgacctacttacaccatgt---ctgtagctatttccctcccttctcttgaact 1231
 Db 304 TGCTGTGACCTNCTTACANCANGTCTGTTNGCTATTATTTCCTGCTGCTGTAAC 363

Qy 1232 ctaagagaaagaaatcctaactgaaatacaaaaaaatacaaaaaaatacaaaaaa 1283
 Db 364 CTCTTAGGAGGAGGATTCCTTAACCTGGGAAATAACCCCAAAAAAANNTTAA 415

RESULT 9
 V30937 standard; DNA; 497 BP.
 ID V30937;
 AC V30937;
 DE Homo sapiens neurotrophin alpha gene related clone HLTBM08R.
 DT neurotrophin alpha; cell proliferation; differentiation; migration;
 KW cytotoxicity; cell death; treatment; tumour; infection; inflammation;
 KW wound healing; immunodeficiency; autoimmune disease; graft rejection;
 KW fibrotic disorder; haematopoiesis; sepsis; shock; malaria; HIV; AIDS;
 KW acquired immune deficiency syndrome; rheumatoid arthritis; silicosis;
 KW cachexia; detection; diagnosis; drug screening; HLTBM08R; ss.
 OS Homo sapiens.


```

PN W09800544-A2.
PD 08-JUN-1998.
PF 27-JUN-1997; U11341.
PR 28-JUN-1996; US-021874.
PA (ACTI-) ACTIVATED CELL THERAPY INC.
PI Engleman EG, Laus R, Ruegg CL;
DR WPI: 98-086969/08.
DR P-PSDB; W45489.
PT New isolated DNA encoding growth arresting protein expressed in B
PT cells - and related vectors, transformed cells and proteins, used
PT for, e.g. gene therapy of leukaemia or diagnosis of cancer and
PT immune status
PS Claim 4; Fig 1; 26pp; English.
CC This cDNA clone, the coding region of which is claimed, codes for
CC a growth arrest protein (see W46489). It corresponds to the B4B
CC gene, expression of which results in inhibition of cellular
CC proliferation. Human peripheral blood mononuclear cells were
CC fractionated to recover high, intermediate and low density
CC fractions. cDNA libraries from all 3 fractions were subjected to
CC a differential display PCR. Bands unique to the intermediate
CC density sample were reamplified, cloned and sequenced. A 101 bp
CC insert was used to screen a cDNA library from intermediate
CC density calls to isolate the 2.7 kb cDNA clone. This insert was
CC used to isolate a genomic fragment localised to the B4B gene to
CC chromosome 20q12-q13.1. The isolated clone can be used to express
CC the growth arrest protein in host cells. Introduction of a
CC functional B4B allele (by gene therapy) into a cell with a
CC defective allele may be used to prevent development of pro-B-cell
CC leukaemia or other malignancies. Introduction of the isolated DNA
CC in a vector that includes a suicide inhibition cassette can be used
CC to arrest expansion of cells, in vivo or in vitro, e.g. useful in
CC skin cell grafts and in lymphocytes containing a heterologous
CC cytokine gene for in vivo expression.
SQ Sequence 2705 Bp; 729 A; 584 C; 627 G; 765 T;

Query Match 3.5%; Score 44.6; DB 1; Length 2705;
Best Local Similarity 65.7%; Pred. No. 0.035;
Matches 65; Conservative 0; Mismatches 34; Indels 0; Gaps 0;

QY 1186 tactacaccatgtctgtagctatttcctccctctctctacacctgaagaagaa 1245
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2600 TTCATATACATAATGTAAATATTTTGTATACAAATGTTTATAACTCTAGGATATAAA 2659
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1246 tctaactgaaatacaaaaaaataaaaaaataaaaaa 1284
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2660 ACAGATCTGATTCACCAAAAAAAAAAAAAAAAAAAAA 2698
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 12
V52934
ID V52934 standard; cDNA; 2639 BP.
AC V52934;
DT 21-DEC-1998 (first entry)
DE Pig transforming growth factor-beta 3 cDNA.
KW Transforming growth factor-beta 3; TGF-beta 3; pig; ss.
OS Sus scrofa.
FH Key Location/Qualifiers
FT CDS 127..1497
FT /tag= a
FT /transl_except= (pos:481..483, aa:Met)
FT US5801231-A.
PD 01-SEP-1998.
PF 22-MAR-1985; 715142.
PR 13-MAR-1987; US-025423.
PR 22-MAR-1985; US-715142.
PR 04-AUG-1989; US-389929.
PR 04-MAR-1992; US-845893.
PR 05-NOV-1993; US-147364.
PR 30-MAY-1995; US-454468.
PA (GETH ) GENENTECH INC.
PI Derynck RNA, Goeddel DV;
DR WPI: 98-494840/42.

```

```

DR P-PSDB; W78786.
PT DNA encoding transforming growth factor-beta precursor sequence -
PT useful for analysis to perform manipulations to increase yield of
PT recombinant production of the protein
PS Example 6; Fig 4a-c; 26pp; English.
CC This nucleotide sequence, termed 10+11.3, codes for the porcine
CC transforming growth factor-beta 3 precursor (preTGF-beta 3, see
CC W78786). A porcine ovarian cDNA library was screened using human
CC TGF-beta 1 cDNA (see V52933) as probe. A hybridising clone,
CC designated lambda 11.3, was used to rescreen the library to
CC identify clone lambda 10. The was combined with clone lambda 11.3
CC to provide the 10+11.3 sequence. The invention relates to the
CC recombinant production of TGF-beta. Biologically active TGF-beta
CC is defined as being capable of inducing EGF-potentiated anchorage
CC independent growth of target cell lines and/or growth inhibition of
CC neoplastic cell lines. Nucleic acids encoding TGF-beta have been
CC isolated and cloned into vectors which are replicated in bacteria
CC and expressed in eukaryotic cells. TGF-beta recovered from
CC transformed cells is used in known therapeutic applications.
CC TGF-beta nucleic acids are also useful in diagnosis and
CC identification of TGF-beta clones.
SQ Sequence 2639 Bp; 677 A; 702 C; 697 G; 563 T;

Query Match 3.5%; Score 44.4; DB 1; Length 2639;
Best Local Similarity 53.4%; Pred. No. 0.039;
Matches 93; Conservative 0; Mismatches 81; Indels 0; Gaps 0;

QY 1111 aatacagagaaatgcacaaatcatcctggatggagatgcacatttttgggcatt 1170
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2457 AAGACAAATACAAATATTACTCTCAAAATCTTTGTATAAATAATTTTGGGAATC 2516
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1171 gaaactgtgtgacctactacacatgtctgtactatttctcctctctgtacc 1230
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2517 TTGGATCATCTTCTCTGGAAGATTGTTCTTAAACAATAAGGCCCTATTTC 2576
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

QY 1231 tctaagaagaagaatcctaactgaaatacaaaaaaataaaaaaataaaaaa 1284
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2577 TAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2630
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
V01527
ID V01527 standard; cDNA to mRNA; 2239 BP.
AC V01527;
DT 21-MAY-1998 (first entry)
DE Wheat soluble starch synthase partial cDNA sequence.
KW Starch synthase; wheat; transgenic plant; ss.
OS Triticum aestivum L. cv. Florida.
FH Key Location/Qualifiers
FT CDS 3..2018
FT /tag= a
FT W09745545-A1.
PN 04-DEC-1997.
PD 04-DEC-1997.
PF 28-MAY-1997; E02793.
PR 11-SEP-1996; DE-036917.
PR 29-MAY-1996; DE-021588.
PA (AGRE ) HOECHST-SCHERING AGREVO GMBH.
PI Block M, Loerz H, Luetticke S, Froberg C, Kossmann J,
PI Walter L;
DR WPI: 98-032652/03.
DR P-PSDB; W23937.
PT Nucleic acid encoding starch synthase enzymes from wheat - for
PT transgenic plants that produce modified forms of starch, useful e.g.
PT in foods, or for production of packaging materials and disposable
PT goods
PS Claim 1; Page 47-51; 71pp; English.
CC This near full-length cDNA clone, designated TaSSS, codes for a
CC soluble starch synthase (see W23837) of summer wheat (cv. Florida).
CC It was isolated from a phage cDNA library of 21-day-old wheat
CC caryopses by screening with a PCR fragment derived from rice soluble
CC starch synthase (see also V01529-30). A second clone (see V01528),
CC coding for wheat granule-bound starch synthase (see W23938) is also

```


GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 1999, 19:44:05 ; Search time 1016.5 Seconds
(without alignments)
688.805 Million cell updates/sec

Title: US-09-212-270-7

Perfect score: 459

Sequence: 1 ctggagcgctgcgcgtactg.....ggctgctgcagggccctcgag 459

Scoring table: IDENTITY_NUC

Searched: 2002476 seqs, 762712212 residues

Database : EST.*

- 1: em_est1.*
- 2: em_est2.*
- 3: em_est3.*
- 4: em_est4.*
- 5: em_est5.*
- 6: em_est6.*
- 7: em_est7.*
- 8: em_est8.*
- 9: em_est9.*
- 10: gb_est1.*
- 11: gb_est2.*
- 12: gb_est3.*
- 13: gb_est4.*
- 14: gb_est5.*
- 15: gb_est6.*
- 16: gb_est7.*
- 17: gb_est8.*
- 18: gb_est9.*
- 19: gb_est10.*
- 20: gb_est11.*
- 21: gb_est12.*
- 22: gb_est13.*
- 23: gb_est14.*
- 24: gb_est15.*
- 25: gb_est16.*
- 26: gb_est17.*
- 27: gb_est18.*
- 28: gb_est19.*
- 29: gb_est20.*
- 30: gb_est21.*
- 31: gb_est22.*
- 32: gb_est23.*
- 33: gb_est24.*
- 34: gb_est25.*
- 35: gb_est26.*
- 36: gb_est27.*
- 37: gb_est28.*
- 38: gb_est29.*
- 39: gb_est30.*
- 40: gb_est31.*
- 41: gb_est32.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|--------------------|
| 1 | 260.6 | 56.8 | 428 | 24 | AA155701 zo70e05.i |
| 2 | 242.2 | 52.8 | 283 | 24 | AA155646 zo70e05.s |

| | | | | | | |
|----|------|------|------|----|----------|----------|
| 3 | 183 | 39.9 | 233 | 15 | AA025672 | AA025672 |
| 4 | 179 | 39.0 | 191 | 17 | AA325843 | EST28933 |
| 5 | 149 | 32.5 | 467 | 15 | AA025673 | AA025673 |
| 6 | 119 | 25.9 | 436 | 23 | AI290210 | AI290210 |
| 7 | 82.8 | 18.0 | 461 | 26 | AA740147 | AA740147 |
| 8 | 78 | 17.0 | 344 | 30 | AI185297 | AI185297 |
| 9 | 67.2 | 14.6 | 378 | 24 | AA613366 | AA613366 |
| 10 | 48 | 10.5 | 358 | 23 | AI282851 | AI282851 |
| 11 | 42.2 | 9.2 | 452 | 15 | AA210969 | AA210969 |
| 12 | 41.2 | 9.0 | 219 | 11 | H49658 | H49658 |
| 13 | 41.2 | 9.0 | 425 | 11 | R81613 | R81613 |
| 14 | 41.2 | 9.0 | 192 | 18 | AA533290 | AA533290 |
| 15 | 41.2 | 9.0 | 469 | 27 | AA833793 | AA833793 |
| 16 | 40.8 | 8.9 | 1257 | 13 | W98479 | W98479 |
| 17 | 40 | 8.7 | 468 | 22 | AA606820 | AA606820 |
| 18 | 40 | 8.7 | 972 | 27 | AA877908 | AA877908 |
| 19 | 39.6 | 8.6 | 632 | 27 | AA888854 | AA888854 |
| 20 | 39.4 | 8.6 | 330 | 24 | AA478319 | AA478319 |
| 21 | 39.2 | 8.5 | 842 | 13 | W28814 | W28814 |
| 22 | 39.2 | 8.5 | 1311 | 14 | W15735 | W15735 |
| 23 | 38.8 | 8.5 | 405 | 12 | H61321 | H61321 |
| 24 | 38.8 | 8.5 | 443 | 19 | AA410243 | AA410243 |
| 25 | 38.2 | 8.3 | 522 | 13 | W94813 | W94813 |
| 26 | 38 | 8.3 | 729 | 28 | AA849564 | AA849564 |
| 27 | 37.8 | 8.2 | 459 | 23 | AI285778 | AI285778 |
| 28 | 37.8 | 8.2 | 441 | 27 | AA896269 | AA896269 |
| 29 | 37.6 | 8.2 | 616 | 30 | AI073725 | AI073725 |
| 30 | 37.4 | 8.1 | 841 | 13 | W28832 | W28832 |
| 31 | 37.2 | 8.1 | 258 | 10 | R49803 | R49803 |
| 32 | 37.2 | 8.1 | 797 | 13 | W26739 | W26739 |
| 33 | 37.2 | 8.1 | 764 | 13 | W27515 | W27515 |
| 34 | 37.2 | 8.1 | 560 | 14 | AA032843 | AA032843 |
| 35 | 37.2 | 8.1 | 391 | 11 | HI6143 | HI6143 |
| 36 | 37 | 8.1 | 1433 | 14 | W15921 | W15921 |
| 37 | 36.8 | 8.0 | 625 | 23 | AA990886 | AA990886 |
| 38 | 36.6 | 8.0 | 337 | 20 | AA088733 | AA088733 |
| 39 | 36.2 | 7.9 | 633 | 14 | AA050152 | AA050152 |
| 40 | 36.2 | 7.9 | 426 | 23 | AI199622 | AI199622 |
| 41 | 36.2 | 7.9 | 466 | 30 | AI088747 | AI088747 |
| 42 | 36.2 | 7.9 | 711 | 31 | AI183403 | AI183403 |
| 43 | 36 | 7.8 | 1461 | 14 | W15723 | W15723 |
| 44 | 36 | 7.8 | 778 | 29 | AI046673 | AI046673 |
| 45 | 36 | 7.8 | 778 | 29 | AI046673 | AI046673 |

ALIGNMENTS

| | |
|------------|--|
| RESULT | 1 |
| AA155701 | AA155701 |
| LOCUS | 428 bp mRNA EST 08-NOV-1997 |
| DEFINITION | zo70e05.r1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592256 5', mRNA sequence. |
| ACCESSION | AA155701 |
| NID | GI727317 |
| KEYWORDS | EST. |
| SOURCE | human. |
| ORGANISM | Homo sapiens |
| REFERENCE | Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. |
| AUTHORS | 1 (bases 1 to 428) Hillier, L., Allen, M., Bowles, J., Dubuque, T., Geisel, G., Jost, S., Krizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R. |
| TITLE | WashU-NCI human EST Project |
| JOURNAL | Unpublished (1997) |
| COMMENT | Contact: Wilson RK Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 |

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 518 Std Error: 0.00
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 213.
Location/Qualifiers
1. .428
/organism="Homo sapiens"
/note="Organ: pancreas; Vector: pBluescript SK-; Site: 1: EcoRI; Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Pancreatic adenocarcinoma cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
/db_xref="GDB:4623642"
/db_xref="taxon:9606"
/clone="592256"
/lab_host="SOLR cells (kanamycin resistant)"
/lab_host="SOLR cells (kanamycin resistant)"
76 a 155 c 114 g 76 t 7 others

BASE COUNT 76 a 155 c 114 g 76 t 7 others
ORIGIN

Query Match 56.8%; Score 260.6; DB 24; Length 428;
Best Local Similarity 87.0%; Pred. No. 4.3e-56;
Matches 375; Conservative 0; Mismatches 45; Indels 11; Gaps 8;

QY 2 tggagcgtccgactctgcaacgctctctgcgggagcgtgagggagcagcggtt 61
Db 1 TGGAGCGCTGCCCTACTGCAAGCTCTCTGCGGGGANGCTGAGGAGGACGCGGCT 60

QY 62 gcaacgcaacccacacccgctgcctgcctgcctgcctgcctgcctgcctgcct 121
Db 61 GCACGCCACCCACCAACCCNGTGCCTGCGCGTGCCTGCCTGCCTGCCTGCCT 119

QY 122 tctgttgagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 191
Db 120 TCTGCTTGANACAGCATGCTGCTCCACCTGCTGCGCGGAGTGGCCCGCC- 178

QY 182 gccagacacgcagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 240
Db 179 GCCAGAACACGCGAGTCCAGCGCTGCCCGCCAGGACCTTCTCAGCCAGCAGCTCCAGC 238

QY 241 tcagacagtcgagcagcagcagcagcagcagcagcagcagcagcagcagc 300
Db 239 TCAGAGCAGTG-CAGCCCCACCCGCACTTGCACGCGCTGGCTGGCCCTCAATGTGCCA 297

QY 300 aggtcttctctccatgacacccctgtgacac-agctgcactggttccccctca 355
Db 298 AGGCTCTCTCCATGACACCTGTGCACCAAGTTGCACGTGCTTCCCCCTCAAGCAAC 357

QY 356 ccaggttacagagagctga--ggagtgtgagcgtgcctgcctgcctgcctgcct 413
Db 358 CAAGGGTACCAGGAGCTGAAGGAAGTGTGAGCGTCCGCTCATGACTTTGTGCTTCCA 417

QY 414 ggcacatctcca 424
Db 418 GGACATCTCCA 428

RESULT 2
AA155646/c
LOCUS
DEFINITION zo70e05.s1 Stratagene pancreas (#937208) Homo sapiens cDNA clone 592256 3', mRNA sequence.
ACCESSION AA155646
NID g1727261
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 518 Std Error: 0.00
Seq primer: -28M13 fwd. from Amersham
High quality sequence stop: 249.
Location/Qualifiers
1. .283
/organism="Homo sapiens"
/note="Organ: pancreas; Vector: pBluescript SK-; Site: 1: EcoRI; Site: 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Pancreatic adenocarcinoma cell line. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGGCAGCAG 3' -3' adaptor sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"
/db_xref="GDB:4623642"
/db_xref="taxon:9606"
/clone="592256"
/lab_host="SOLR cells (kanamycin resistant)"
/lab_host="SOLR cells (kanamycin resistant)"
51 a 69 c 111 g 49 t 49 others

BASE COUNT 51 a 69 c 111 g 49 t 49 others
ORIGIN

Query Match 52.8%; Score 242.2; DB 24; Length 283;
Best Local Similarity 96.9%; Pred. No. 1.7e-51;
Matches 277; Conservative 0; Mismatches 6; Indels 3; Gaps 3;

QY 140 cgtgtcacactggtgcggcggtgattgccccgggacccccagcagacagcagtc 199
Db 283 CTGTGTCACCTGGTGGCGCGTGATT-NCCCGGGCACCCCGCCAGCAGACACGAGTGCC 225

QY 200 agcgtgtccccccagcagcagcagcagcagcagcagcagcagcagcagcagc 259
Db 224 AGCCNT-CCCCCAGCAGCAGCTTCTCAGCCAGCAGCTCCAGCTCAGCAGTGCCAGCCCC 166

QY 260 accgcaactgcagcggcctggcctggcctggcctggcctggcctggcctggcct 319
Db 165 ACCGCAACTGCACAGCCCT-GGNTGGCCCTCAATGTGCAGGCTCTTCTCCCATGACA 107

QY 320 cctgtgacacagtcagtcagtcagtcagtcagtcagtcagtcagtcagtcag 379
Db 106 CCTGTGACACAGCTGCATGGCTTCCCTCAGCAGCAGGAGTACAGGAGTGAGGAGT 47

QY 380 gtgagcgtgcccgtcatcgactgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 425
Db 46 GTGAGCGTCCGCTCATCGACTTTTGGGCTTTCAGGACATCTCCAT 1

RESULT 3
AA025672
LOCUS
DEFINITION ze90h09.r1 Soares fetal heart NbHH19W Homo sapiens cDNA clone 366305 5', mRNA sequence.
ACCESSION AA025672
NID g1491065
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata;

REFERENCE 1 (bases 1 to 233)
 AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.
 TITLE WashU-Merck EST Project
 JOURNAL Unpublished (1995)
 COMMENT
 Contact: Wilton RK
 WashU-Merck EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
 Insert Length: 713 Std Error: 0.00
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 211.
 Location/Qualifiers
 1..233
 /organism="Homo sapiens"
 /note="Organ: heart; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTACCAATCTCAAGTGGAGCGCGCATCTTTTCTTTTCTTTT 3'] double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by M. Fatima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."
 /db_xref="taxon:9606"
 /clone="366305"
 /clone_lib="Soares fetal heart NBHL19W"
 /sex="unknown"
 /dev_stage="19 weeks"
 /lab_host="DH10B (ampicillin resistant)"
 <1..>233
 BASE COUNT 40 a 96 c 57 g 35 t 5 others
 ORIGIN
 Query Match 39.9%; Score 183; DB 15; Length 233;
 Best Local Similarity 96.2%; Pred. No. 1.1e-36;
 Matches 227; Conservative 0; Mismatches 5; Indels 4; Gaps 4;
 QY 136 gcatcgtgtccactgtgctggcggtgattgcccccggcaccacacacacgag 195
 Db 1 GCATCGTGTCCACCTGGTGGCGGGGTATTGCCCC-GGNACCCCGAGCAGACACGCGAG 59
 QY 196 tgccagcgtgtcccccagaccac-ttctcagcagcagctccagctcagcagtgcca 254
 Db 60 -GCCAGCGGTGCCCGCCCGAGGACCTTTCTCAGCAGCAGCTCCAGCTCAGACAGTGCCA 118
 QY 255 gccccacgcgaactgcagcgccgtggcgctccctcaatgtgcaggctcttctccca 314
 Db 119 GCCCCACGGCAACTGCACGGNCCI-GGNNTGGCCCTCAATGTGCAGGCTCTTCTCTCCA 177
 QY 315 tgacacctgtgaccagctgactggtctccctccctcagcaccaggggtaccagag 370
 Db 178 TGACACCGTGTGCACCAAGCTCAGCTGGGCTTCCCTCTCAGCACCGAGGTACCANGAG 233
 RESULT 4
 AA325843 191 bp mRNA EST 20-APR-1997
 LOCUS AA325843
 DEFINITION EST28933 Cerebellum II Homo sapiens cDNA 5' end, mRNA sequence.
 ACCESSION AA325843

NID
 KEYWORDS
 SOURCE
 ORGANISM

gi1978108
 EST.
 human.
 Homo sapiens
 Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Euthera; Primates; Catarrhini; Homiidae;
 Homo.

REFERENCE
 AUTHORS

1 (bases 1 to 191)
 Adams, M.D., Kerlavage, A.R., Fleischmann, R.D., Fuldner, R.A., Bult, C.J., Lee, N.H., Kirkness, E.F., Weinstock, K.G., Gocayne, J.D., White, O., Sutton, G., Blake, J.A., Brandon, R.C., Man-Wai, C., Clayton, R.A., Cline, T.R., Cotton, M.D., Earle-Hughes, J., Fine, L.D., Fitzgerald, L.M., Fitzhugh, W.M., Fritchman, J.L., Geoghagen, N.S., Glodek, A., Gnehm, C.L., Hanna, M.C., Hedblom, E., Hinkle, P.S., Jr., Kelley, J.M., Kelley, J.C., Liu, L.-I., Marmaros, S.M., Merrick, J.M., Moreno-Palancas, R.F., McDonald, L.A., Nguyen, D.T., Pelligrino, S.M., Phillips, C.A., Ryder, S.E., Scott, J.L., Saudek, D.M., Shirley, R., Small, K.V., Spriggs, T.A., Utterback, T.R., Weidman, J.F., Li, Y., Bednarek, D.P., Cao, L., Cepeda, M.A., Coleman, T.A., Collins, E.J., Dimke, D., Feng, D.-F., Ferrie, A., Fischer, C., Hastings, G.A., He, W.W., Hu, J.S., Greene, J.M., Gruber, J., Hudson, P., Kim, A.K., Kozak, D.L., Kunsch, C., Hungjun, J., Li, H., Meissner, P.S., Olsen, H., Raymond, L., Wei, Y.F., Wing, J., Xu, C., Yu, G.L., Ruben, S.M., Dillion, P.J., Fannon, M.R., Rosen, C.A., Haseltine, W.A., Fields, C., Fraser, C.M., and Venter, J.C.
 Initial assessment of human gene diversity and expression patterns based upon 83 million nucleotides of cDNA sequence
 Nature 377 (6547 Suppl), 3-174 (1995)
 JOURNAL
 MEDLINE
 COMMENT
 Other_ESTs: THC162279
 Contact: Kerlavage, AR
 Bioinformatics
 The Institute for Genomic Research
 9712 Medical Center Drive, Rockville, MD 20850 USA
 Tel: 3018699056
 Fax: 3018699423
 Email: arkerlav@tigr.org
 For clone availability, additional sequence and expression information related to this EST, please check the TIGR Human Gene Index (<http://www.tigr.org/tdb/hgi/hgi.html>)
 Seq primer: M13 Reverse.
 Location/Qualifiers
 1..191
 /organism="Homo sapiens"
 /note="Organ: brain; Vector: pBluescript SK-; Site_1: EcoRI; Site_2: XhoI"
 /db_xref="ATCC (inhos):126377"
 /db_xref="taxon:9606"
 /clone_lib="Cerebellum II"
 /tissue_type="cerebellum"
 /dev_stage="adult"
 <1..>191

FEATURES
 SOURCE

BASE COUNT 33 a 66 c 49 g 39 t 4 others
 ORIGIN
 mRNA
 Query Match 39.0%; Score 179; DB 17; Length 191;
 Best Local Similarity 95.3%; Pred. No. 1.1e-35;
 Matches 182; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 QY 262 cgcaactgcacggccctggcgccctcaatgtgcaggctcttctccatgcacac 321
 Db 1 CGCAACTGCACGGCCCTGGGAGCTGCCCTCAATGTGCAGGNTCTTCTCCCATGACACC 60
 QY 322 ctgtgcaccagctgactggtctccctcagcaccaggtaccagagctgagatgt 381
 Db 61 CTGTGCACCAAGCTGCATGGCTTCCCTCAGCACCAGGTACCANGAGCTGAGGATGT 120
 QY 382 gacgctgccgtcatcactgtgtgttccatgcacacatctccatcaagaggtcagcgg 441
 Db 121 GAGCNTCCGCTCATCGACTTTTGGCTTTCAGGACATCTCCATCAGAGGCTGCACGG 180
 QY 442 ctgctgcaggc 452


```

Db 181 CTGCTCANGCC 191
|||||
RESULT 5
AA025673 467 bp mRNA EST 01-FEB-1997
LOCUS ze90h09.s1 Soares fetal heart NBHH19W Homo sapiens cDNA clone
DEFINITION 36305 3', mRNA sequence.
ACCESSION AA025673
NID g1491066
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondria; eukaryotes; Metazoa; Chordata;
Vertebrata; Euthera; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 467)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,
Waterston, R., Williamson, A., Wohlmann, P., and Willson, R.
WASHU-Merck EST Project
WASHU-Merck EST Project
Contact: Wilson RK
WASHU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 713 Std Error: 0.00
Seq primer: -40M13 fwd. from Amersham
High quality sequence stop: 373.
FEATURES
Source
1..467
/organism="Homo sapiens"
/note="Organ: heart; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not 1; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAAGTGGGCGCGCATCTTTTCTTTTCTTTT 3'], 1st
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Fatima Bonaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBHH19W."
/db_xref="taxon:9606"
/clone="366305"
/clone_lib="Soares fetal heart NBHH19W"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>467)
BASE COUNT 94 a 135 c 134 g 99 t 5 others
ORIGIN
mRNA
Query Match 32.5%; Score 149; DB 15; Length 467;
Best Local Similarity 95.6%; Pred. No. 3.4e-28;
Matches 152; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy 300 aggccttcctccatgacacccctgtgaccagctgacctgctccctccctccagcaccag 359
Db 467 AGNCCTCTTCTTCTCCCATGACACCTGTGACACCGCTGCACTGGCTTCCCTCCAGCACCAG 408
Qy 360 ggtaccagagctgaggagctgagcgtgccctcatcactgtgctcttcctccagacat 419
Db 407 GNTACCAGAGCTGAGGAGTGTGAGCGTCCGCTCATCGACTTGTGGCTTCTCCAGGACAT 348

Qy 420 ctccatcaagagctgcagcggctgctgcagcgcctcga 458
|||||
Db 347 CTCCATCAAGAGGCTCAGCGGCTGCTGCAAGCCCTCGA 309

RESULT 6
AI290210/c 436 bp mRNA EST 30-NOV-1998
LOCUS q179g12.x1 Soares_NhMPu_S1 Homo sapiens cDNA clone IMAGE:1878598
DEFINITION 3', mRNA sequence.
ACCESSION AI290210
NID g3931864
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Euthera;
Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 436)
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40Up from Gibco
High quality sequence stop: 388.
FEATURES
Source
1..436
/organism="Homo sapiens"
/note="Organ: mixed (see below); Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not 1;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NbHPU, and fetal heart NBHH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."
/db_xref="taxon:9606"
/clone="IMAGE:1878598"
/clone_lib="Soares_NhMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/lab_a 97 a 136 c 124 g 79 t
BASE COUNT
ORIGIN
Query Match 25.9%; Score 119; DB 23; Length 436;
Best Local Similarity 100.0%; Pred. No. 1.1e-20;
Matches 119; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 341 gcttcctccctcagaccaggggtaccagagctgagagctgagcgtccctccctcagcact 400
Db 436 GCTTCCCTCCCTCAGCACCAGGCTACCCAGGCTGAGGAGTGTGAGCGTCCGCTCATCGACT 377
Qy 401 ttgtggctttccagacatctccatcaagagcgtgcagcgtctgagcgtccctcag 459
Db 376 TTGTGGCTTTCAGGACATCTCCATCAAGAGGCTCAGCGGCTGCTGCAGGCCCTCGAG 318

RESULT 7
AA740147/c 461 bp mRNA EST 08-FEB-1998
LOCUS ob26a08.s1 NC1_CGAP_Kid5 Homo sapiens cDNA clone IMAGE:1324790,
DEFINITION mRNA sequence.

```


Email: Robert_Strausberg@nih.gov
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 342.

FEATURES
source
1. .344
/organism="Homo sapiens"
/note="Organ: lung; Vector: pT7T3D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-TGTTACCAATCTGAAGTGGAGCGCCCAATTTTTTTTTTTT-3'] double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT7T3 vector (Pharmacia). Library went through one round of normalization to a Cot = 5. Library constructed by Bento Soares and M.Fatima Bonaldo. This library was constructed from the same fetus as the fetal heart library, Soares fetal heart NBHL19W."
/db_xref="taxon:9606"
/clone="IMAGE:1741058"
/clone_lib="Soares fetal_lung_NBHL19W"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
BASE COUNT 72 a 115 c 106 g 51 t
ORIGIN

Query Match 17.0%; Score 78; DB 30; Length 344;
Best Local Similarity 100.0%; Pred. NO. 1.9e-10;
Matches 78; Conservative 0; Mismatches 0; Indels 0; Gaps 0

Qy 382 gagcgctgcctcatcgacttggctttccaggacatctccatcaaggctcgagcg 441
|||||
Db 344 GAGCGTCCGTCATCGACTTGTGGCTTCCAGGACATCTCCATCAAGGCTGCAGCG 285
|||||

Qy 442 ctgctgcaggcctctcgag 459
|||||
Db 284 CTGCTGCAGGCCCTCGAG 267

RESULT 9
AA613366/c
LOCUS
DEFINITION
ACCESSION
NID
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA613366 378 bp mRNA EST 16-OCT-1997
nc31a08.s1 NCI_CGAP_Co10 Homo sapiens cDNA clone IMAGE:1145462,
mRNA sequence.
AA613366
92464404
EST.
human.
Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 378)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Ilan Kirsch, M.D., Michael R. Emmert-Buck,
M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 978 Std Error: 0.00

AA740147
92778739
EST.
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AA740147
92778739
EST.
human.
Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 461)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
unknown library type
Insert Length: 1402 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham.
Location/Qualifiers
1. .461
/organism="Homo sapiens"
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5'-AATCGAAGATTCGGCGCGCAATTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1324790"
/clone_lib="NCI_CGAP_Kid5"
/tissue_type="2 pooled tumors (clear cell type)"
/lab_host="DH10B"
BASE COUNT 84 a 148 c 158 g 71 t
ORIGIN

Query Match 18.0%; Score 82.8; DB 26; Length 461;
Best Local Similarity 92.6%; Pred. NO. 1.2e-11;
Matches 87; Conservative 0; Mismatches 7; Indels 0; Gaps 0

Qy 365 caggagctgaggagtgagcgtgcctcatcgacttggctttccaggacatctcca 424
|||||
Db 371 CAGGAGCTGAGGAGTGAGCGTGGCGTCATCGACTTGTGGCTTCCAGGACATCTCCA 312
|||||

Qy 425 tcaagagctgcaggcgtgctctcgagcctctga 458
|||||
Db 311 TCAAGAGCGTGAAGCGGCTGCTGCCAGCGCTCGA 278
|||||

RESULT 8
AI185297/c
LOCUS
DEFINITION
ACCESSION
NID
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AI185297 344 bp mRNA EST 09-OCT-1998
ge36c02.s1 Soares_fetal_lung_NBHL19W Homo sapiens cDNA clone
IMAGE:1741058 3', mRNA sequence.
AI185297
93735935
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 344)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550

```

Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 367.
Location/Qualifiers
1..378
/organism="Homo sapiens"
/note="Organ: colon; Vector: pT7T3D-Pac (Pharmacia) with a
modified polylinker; 1st strand cDNA was prepared from
RER+ colon tumor, and was then primed with a Not I -
oligo(dT) primer. Double-stranded cDNA was ligated to Eco
RI adaptors (Pharmacia), digested with Not I and cloned
into the Not I and Eco RI sites of the modified pT7T3
vector. Library is normalized. Library was constructed by
Bento Soares and M. Fatima Bonaldo (N-Soares4)."
/db_xref="taxon:9606"
/clone="IMAGE:1145462"
/clone_lib="NCI_CGAP_Col10"
/tissue_type="colon tumor RER+"
/lab_host="DH10B"

BASE COUNT      90 a 117 c 109 g 62 t
ORIGIN

Query Match      14.6%; Score 67.2; DB 24; Length 378;
Best Local Similarity 95.8%; Pred. No. 9.7e-08;
Matches 69; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 388 gcgcgcacgtttgtgcttcacaggacatccatcaagaggtgcgcggtgctg 447
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 GCGGTCATCGACTTTGTGGCTTTCCAGGACATCTCCATCAAGAGGCTCGAGGGCTGCTG 319

QY 448 caggccctcgag 459
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 CATGCCCTCGAG 307

RESULT 10
LOCUS      A1282851/c 358 bp mRNA EST 23-NOV-1998
DEFINITION qt84e09.x1 NCI_CGAP_Col14 Homo sapiens cDNA clone IMAGE:1961992 3',
            mRNA sequence.
ACCESSION  A1282851
NID        G3921084
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 358)
AUTHORS    NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
JOURNAL    Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Tel: (301) 496-1550
            Email: Robert.Strausberg@nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 323.
Location/Qualifiers
1..358
/organism="Homo sapiens"
/note="Organ: colon; Vector: pCMV-SPORT6; Site.1: SalI;
Site.2: NotI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.7 kb. Life Technologies catalog #:

FEATURES
source

```

```

11531-019"
/db_xref="taxon:9606"
/clone="IMAGE:1961992"
/clone_lib="NCI_CGAP_Col14"
/tissue_type="moderately-differentiated adenocarcinoma"
/lab_host="DH10B"

BASE COUNT      81 a 114 c 103 g 60 t
ORIGIN

Query Match      10.5%; Score 48; DB 23; Length 358;
Best Local Similarity 100.0%; Pred. No. 0.0061;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 412 caggacatccatcaagaggtgcgcggtgctgcagccctcgag 459
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 CAGGACATCTCCATCAAGAGGCTCGAGGGCTGCTGCAGGCCCTCGAG 311

RESULT 11
LOCUS      AA210969 452 bp mRNA EST 31-JAN-1997
DEFINITION zq88c08.r1 Stratagene hNT neuron (#937233) Homo sapiens cDNA clone
            649070 5' similar to gb:J03827 Y BOX BINDING PROTEIN-1 (HUMAN);,
            mRNA sequence.
ACCESSION  AA210969
NID        G1809632
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
            Vertebrata; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 452)
AUTHORS    Hillier, L., Clark, N., Dubouque, T., Elliston, K., Hawkins, M.,
            Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
            Parsons, J., Rifkin, L., Rohlfing, T., Tan, F., Trevaskis, E.,
            Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R.
            WashU-Merck EST Project
            Unpublished (1995)
TITLE      Contact: Wilson RK
            WashU-Merck EST Project
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            This clone is available royalty-free through LLNL; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Seq primer: -28ml3 rev1 ET from Amersham
            High quality sequence stop: 331.
            Location/Qualifiers
            1..452
            /organism="Homo sapiens"
            /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
            XhoI; Cloned unidirectionally. Primer: Oligo dt.
            Differentiated, post mitotic hNT neurons. Average insert
            size: 1.5 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5',
            GAATTCGGCAGGAG 3', -3' adaptor sequence: 5',
            CTCGAGTTTCTTTTCTTTTCTTTT 3'."
            /db_xref="taxon:9606"
            /clone="649070"
            /clone_lib="Stratagene hNT neuron (#937233)"
            /dev_stage="hNT neurons"
            /lab_host="SOLR (kanamycin resistant)"
            <1..>452

BASE COUNT      113 a 142 c 128 g 59 t 10 others
ORIGIN

Query Match      9.2%; Score 42.2; DB 15; Length 452;
Best Local Similarity 52.8%; Pred. No. 0.17;
Matches 86; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

```

Best Local Similarity 34.8%; Pred. NO. 0.31;
Matches 102; Conservative 0; Mismatches 83; Indels 1; Gaps 1;

This Page Blank (uspto)


```

Db 513 TGCAGAAATGTCAGATGGTTCTTCTCAGGTGAGACTTCAGAAAGCACCTGTATA 572
QY 256 cccacgcgaactgcagcgccctggccctgacctcaatgtgccaggctcttcccat 315
Db 573 AAMACAGAACTGCAGACACTTGGCCCTCCTGCTAATTCAGAAAGGAATGCAACACAT 632
QY 316 gacacctgtgcaccagctgca 337
Db 633 GACACGCTGTGTTCCGGARACA 654

RESULT 8
T33180
ID T33180 standard; DNA; 564 BP.
AC T33180;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CSph, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT signal_peptide 1..63 /*tag= a
FT mat_peptide 64..561 /*tag= b
FT /*product= OCIF-CSph
PN W09626217-Al.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Iano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB; R99950.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 84; Page 149-150; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-CSph in which amino acids 167-380 of the mature
CC protein have been deleted. These amino acid changes have been caused
CC by the introduction of a restriction site. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 564 BP; 159 A; 149 C; 132 G; 124 T;

```

```

Query Match 15.0%; Score 69; DB 1; Length 564;
Best Local Similarity 50.8%; Pred. No. 1.8e-07;
Matches 165; Conservative 0; Mismatches 160; Indels 0; Gaps 0;

QY 10 tgccgtactgcaacgtctctgtcgaggagcgtgagaggagcagcggttgcacgcc 69
Db 238 TGTCTATACTGCAGCCCGTGTGCAAGGAGCTGCAGTACGTCAAGCAGGAGTCAATCGC 297
QY 70 acccaacaacgtgcctgcgcgtgcgcacacggcttcttcgcgcacgtggttctgttg 129
Db 298 ACCACACACCGCGTGTGCGAATGCAAGGAGCGCGTACCTTGAGATAGAGTTCTGCTTG 357
QY 130 gagcagcatgctccacctggtgcgcgtgattgccccgggacccccagccagaac 189
Db 358 AAMCATAGAGAGCTGCCCTCCTGGATTGGAGTGGTGCAGAGCTGGAACCCAGAGCGAAAT 417
QY 190 acgcagtgccagcgtgccccccagccaccttcttcagccagcagctccagctcagagcag 249

```

```

Db 418 ACAGTTTGCAGAAAGATGTCAGATGGTTCTTCTCAATGAGACGTCAATCAAGCACCC 477
QY 250 tgccagcccccacgcgaactgcagcgccctggccctgacctcaatgtgccaggctcttc 309
Db 478 TGTAGAAAAACACACAATTCAGTGTCTTTGGTCTCCTGCTAACTCAGAAAGGAATGCA 537
QY 310 tcccatgacacccctgtgcaccagct 334
Db 538 ACACAGCAGACACATATGTTCCGGCT 562

RESULT 9
T33179
ID T33179 standard; DNA; 966 BP.
AC T33179;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CSph, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
OS Synthetic.
FH Key Location/Qualifiers
FT signal_peptide 1..63 /*tag= a
FT mat_peptide 64..963 /*tag= b
FT /*product= OCIF-CSph
PN W09626217-Al.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Iano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB; R99949.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 81; Page 149; 183pp; Japanese.
CC This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-CSph in which amino acids 298-380 of the mature
CC protein have been deleted and replaced by Ser-Leu-Asp. These amino
CC acid changes have been caused by the introduction of a restriction
CC site. The OCIF of the invention has a molecular weight by SDS-PAGE of
CC 60 kD under reducing conditions and 120 kD under non-reducing
CC conditions. The protein is adsorbed onto cation-exchangers or heparin
CC and its activity is lowered after 10 mins at 70 deg.C or 30 mins at 56
CC deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful in the
CC control of bone resorption and therefore in the treatment and
CC prevention of disorders of bone resorption, e.g. osteoporosis.
SQ Sequence 966 BP; 301 A; 228 C; 226 G; 211 T;

```

```

Query Match 14.9%; Score 68.6; DB 1; Length 966;
Best Local Similarity 50.5%; Pred. No. 2.4e-07;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 10 tgccgtactgcaacgtctctgtcgaggagcgtgagaggagcagcggttgcacgcc 69
Db 238 TGTCTATACTGCAGCCCGTGTGCAAGGAGCTGCAGTACGTCAAGCAGGAGTCAATCGC 297
QY 70 acccaacaacgtgcctgcgcgtgcgcacacggcttcttcgcgcacgtggttctgttg 129
Db 298 ACCACACACCGCGTGTGCGAATGCAAGGAGCGCGTACCTTGAGATAGAGTTCTGCTTG 357
QY 130 gagcagcatgctccacctggtgcgcgtgattgccccgggacccccagccagaac 189
Db 358 AAMCATAGAGAGCTGCCCTCCTGGATTGGAGTGGTGCAGAGCTGGAACCCAGAGCGAAAT 417
QY 190 acgcagtgccagcgtgccccccagccaccttcttcagccagcagctccagctcagagcag 249
Db 418 ACAGTTTGCAGAAAGATGTCAGATGGTTCTTCTCAANTGAGACGTCAATCAAGCACCC 477

```


[illegible]

Db 415 ACACAGCAACAATATGTTCCGGAAACAGTG 445

RESULT 13
T33170

ID T33170 standard; DNA; 981 BP.
AC T33170:
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DDD1, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT signal_peptide 1..63 /*tag= a
FT mat_peptide 64..978 /*tag= b
FT /product= OCIF-DDD1
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB: R99940.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 54; Page 141-142; 183pp; Japanese.
CS This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-DDD1 in which amino acids 178-252 of the mature
CC protein have been deleted. The OCIF of the invention has a molecular
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30
CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful
CC in the control of bone resorption and therefore in the treatment and
CC prevention of disorders of bone resorption, e.g. osteoporosis.
SQ Sequence 981 BP; 312 A; 238 C; 218 G; 213 T;

Query Match 14.9%; Score 68.6; DB 1; Length 981;
Best Local Similarity 50.5%; Pred. No. 2.4e-07;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 10 tgccgctactgaacgtcctctgcggaggcgtaggagggcacgggcttgcacgcc 69
||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 TGCTATACTGCAGCCCCGTGTGCANGAGCGTCGAGTAGCTCAAGCAGGAGTGAATCGC 297
||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 acccaaacacctgcctgcgcgtgcgcaccggcttcctgcgcacgtgtttctgttg 129
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 ACCCACAAACCGGTGTGCGAATGCAAGGAGCGCTACCTTGAGATAGAGTTCTGCTTG 357
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 gagcagcatcgtgtccaacctggtgcggcggtgattccccgggcccacccagcagaac 189
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 AAACATAGAGAGTGCCTCTCTGGATTGGAGTGTGTAAGTGAAGTGAAGTGAAGTGAAGT 417
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 acgcagtgcacgcgtgcgcacccagcaccttcctcagcagcagctccagctcagagcag 249
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 ACAGTTTGCAAAGAATGCCAGATGGGTTCCTCTCAATGNAGAGCTCATCTAAGACACC 477
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 250 tgcacgccccaccgcaactgcacggccctggccctggccctcaatgtgcaggctcttc 309
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 478 TGTAGAAAACACACAAATTCAGTGTCTTTGGTCTCTCTGCTCACTCAGAAAGGAATGCA 537
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 tcccatgacacctgtgcaccagctgcactg 340
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 ACACAGCAACAATATGTTCCGGAAACAGTG 568

Db 310 tcccatgacacctgtgcaccagctgcactg 340
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 538 ACACAGCAACAATATGTTCCGGAAACAGTG 568

RESULT 12
T33166

ID T33166 standard; DNA; 1083 BP.
AC T33166;
DT 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCR1, coding sequence.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW Osteoporosis; ss.
OS Synthetic.
FH Key Location/Qualifiers
FT signal_peptide 1..63 /*tag= a
FT mat_peptide 64..1080 /*tag= b
FT /product= OCIF-DCR1
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR P-PSDB: R99936.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 42; Page 137-138; 183pp; Japanese.
CS This sequence encodes a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence encodes OCIF-DCR1 in which amino acids 2-42 of the mature
CC protein have been deleted. The OCIF of the invention has a molecular
CC weight by SDS-PAGE of 60 kD under reducing conditions and 120 kD under
CC non-reducing conditions. The protein is adsorbed onto cation-exchangers
CC or heparin and its activity is lowered after 10 mins at 70 deg.C or 30
CC mins at 56 deg.C, and is lost after 10 mins at 90 deg.C. OCIF is useful
CC in the control of bone resorption and therefore in the treatment and
CC prevention of disorders of bone resorption, e.g. osteoporosis.
SQ Sequence 1083 BP; 352 A; 250 C; 246 G; 235 T;

Query Match 14.9%; Score 68.6; DB 1; Length 1083;
Best Local Similarity 50.5%; Pred. No. 2.4e-07;
Matches 167; Conservative 0; Mismatches 164; Indels 0; Gaps 0;

QY 10 tgccgctactgaacgtcctctgcggaggcgtaggagggcacgggcttgcacgcc 69
||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 115 TGCTATACTGCAGCCCCGTGTGCAAGAGCTCGAGTAGCTCAAGCAGGAGTGAATCGC 174
||| ||||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 70 acccaaacacctgcctgcgcgtgcgcaccggcttcctgcgcacgtgtttctgttg 129
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 175 ACCCACAAACCGGTGTGCGAATGCAAGGAGCGCTACCTTGAGATAGAGTTCTGCTTG 234
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 130 gagcagcatcgttgtccaacctggtgcggcggtgattccccgggcccacccagcagaac 189
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 235 AAACATAGAGAGTGCCTCTCTGGATTGGAGTGTGTAAGTGAAGTGAAGTGAAGTGAAGT 294
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 190 acgcagtgcacgcgtgcgcacccagcaccttcctcagcagcagctccagctcagagcag 249
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 295 ACAGTTTGCAAAGAATGCCAGATGGGTTCCTCTCAATGNAGAGCTCATCTAAGACACC 354
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 250 tgcacgccccaccgcaactgcacggccctggccctggccctcaatgtgcaggctcttc 309
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 355 TGTAGAAAACACACAAATTCAGTGTCTTTGGTCTCTCTGCTCACTCAGAAAGGAATGCA 414
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 310 tcccatgacacctgtgcaccagctgcactg 340
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 03:30:50 ; Search time 2254.51 Seconds
(without alignments)
728.360 Million cell updates/sec

Title: US-09-212-270-7

Perfect score: 459

Sequence: 1 ctggagcgctgcgcctactg.....ggctgctgcagggccctcgag 459

Scoring table: IDENTITY_NUC

Searched: 808301 seqs, 1788773984 residues

Database : GenEmbl.*

- 1: gb_ba1.*
- 2: gb_ba2.*
- 3: gb_in.*
- 4: gb_on.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl1.*
- 9: gb_pl2.*
- 10: gb_pr1.*
- 11: gb_pr2.*
- 12: gb_pr3.*
- 13: gb_ro.*
- 14: gb_st.*
- 15: gb_sy.*
- 16: gb_un.*
- 17: gb_v1.*
- 18: gb_htg.*
- 19: em_ba.*
- 20: em_fun.*
- 21: em_hum1.*
- 22: em_hum2.*
- 23: em_in.*
- 24: em_on.*
- 25: em_or.*
- 26: em_ov.*
- 27: em_pat.*
- 28: em_ph.*
- 29: em_pl.*
- 30: em_ro.*
- 31: em_sy.*
- 32: em_un.*
- 33: em_v1.*
- 34: em_htg.*
- 35: em_sts.*
- 36: gb_ba1.*
- 37: gb_ba2.*
- 38: gb_pl1.*
- 39: gb_pl2.*
- 40: gb_pr1.*
- 41: gb_pr2.*
- 42: gb_pr3.*
- 43: gb_sts.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|-------|--------|-------|-------|-------------|
| ----- | ----- | ----- | ----- | ----- | ----- | ----- | ----- |

| | | | | | | |
|------|------|------|-------|----|------------|--------------------|
| 1 | 83.2 | 18.1 | 2432 | 13 | RNU94330 | U94330 Rattus norv |
| 2 | 75.6 | 16.5 | 1325 | 13 | MMU94331 | U94331 Mus musculu |
| 3 | 74 | 16.1 | 2818 | 13 | AB013898 | AB013898 Mus muscu |
| 4 | 68.6 | 14.9 | 1206 | 11 | AB002146 | AB002146 Homo sapi |
| 5 | 68.6 | 14.9 | 1356 | 11 | HSU94332 | U94332 Human osteo |
| 6 | 68.6 | 14.9 | 1206 | 41 | AB002146 | AB002146 Homo sapi |
| 7 | 68.6 | 14.9 | 1356 | 41 | HSU94332 | U94332 Human osteo |
| 8 | 52.8 | 11.5 | 1388 | 13 | MMTNFR2A | X76401 M.musculus |
| 9 | 51.2 | 11.2 | 3795 | 13 | MUSTNFR1 | M59378 Murine tumo |
| 10 | 51.2 | 11.2 | 1505 | 13 | MUSTNFR2 | M60469 Mouse tumor |
| 11 | 48.6 | 10.6 | 2116 | 13 | AF019046 | AF019046 Mus muscu |
| 12 | 47.2 | 10.3 | 1968 | 12 | AF068868 | AF068868 Homo sapi |
| 13 | 47.2 | 10.3 | 1968 | 42 | AF068868 | AF068868 Homo sapi |
| 14 | 45 | 9.8 | 2339 | 6 | A26415 | A26415 cDNA fragme |
| 15 | 44.4 | 9.7 | 1600 | 6 | A43530 | A43530 Sequence 1 |
| 16 | 44.4 | 9.7 | 3386 | 11 | HUMNGFR | M14764 Human nerve |
| 17 | 44.4 | 9.7 | 3386 | 41 | HUMNGFR | M14764 Human nerve |
| 18 | 43.8 | 9.5 | 9900 | 11 | AB008821S2 | AB008822 Homo sapi |
| 19 | 43.8 | 9.5 | 9900 | 41 | AB008821S2 | AB008822 Homo sapi |
| 20 | 43.6 | 9.5 | 1080 | 10 | HUMTNFRA | M35857 Human tumor |
| 21 | 43.6 | 9.5 | 1080 | 40 | HUMTNFRA | M35857 Human tumor |
| 22 | 43.4 | 9.5 | 1641 | 6 | I36196 | I36196 Sequence 1 |
| 23 | 43.4 | 9.5 | 1557 | 6 | I36197 | I36197 Sequence 3 |
| 24 | 43.4 | 9.5 | 691 | 6 | I36350 | I36350 Sequence 12 |
| 25 | 43.4 | 9.5 | 7218 | 6 | I66494 | I66494 Sequence 14 |
| 26 | 43.4 | 9.5 | 3683 | 10 | HUMNFR | M32315 Human tumor |
| 27 | 43.4 | 9.5 | 2394 | 10 | HUMTNFRII | M55994 Human tumor |
| 28 | 43.4 | 9.5 | 3492 | 10 | S63368 | S63368 tumor necro |
| 29 | 43.4 | 9.5 | 380 | 13 | AB013899S2 | AB013900 Mus muscu |
| 30 | 43.4 | 9.5 | 3683 | 40 | HUMNFR | M32315 Human tumor |
| 31 | 43.4 | 9.5 | 2394 | 40 | HUMTNFRII | M55994 Human tumor |
| 32 | 43.4 | 9.5 | 3492 | 40 | S63368 | S63368 tumor necro |
| 33 | 43.4 | 9.5 | 3380 | 43 | G26865 | G26865 human STS S |
| c 34 | 43.2 | 9.4 | 7412 | 17 | PVUL50S | X87246 Pseudorabie |
| 35 | 43 | 9.4 | 4080 | 17 | ADVR | X72087 Aujeszky's |
| 36 | 43 | 9.4 | 5657 | 17 | PVPCA | X80797 Pseudorabie |
| c 37 | 42.8 | 9.3 | 4000 | 13 | MUSPOUDOMA | M88299 Mouse brain |
| 38 | 42.6 | 9.3 | 5823 | 3 | DROZFH1 | M63449 D.melanogas |
| 39 | 42.6 | 9.3 | 2168 | 10 | HUMPSDBPB | L37516 Homo sapien |
| 40 | 42.6 | 9.3 | 2168 | 40 | HUMPSDBPB | L37516 Homo sapien |
| c 41 | 42.4 | 9.2 | 2461 | 6 | I23785 | I23785 Sequence 1 |
| c 42 | 42.4 | 9.2 | 2461 | 10 | HUMGAS | L13720 Homo sapien |
| c 43 | 42.4 | 9.2 | 2461 | 40 | HUMGAS | L13720 Homo sapien |
| c 44 | 42 | 9.2 | 3060 | 3 | AF027735 | AF027735 Nephila c |
| c 45 | 41.6 | 9.1 | 28826 | 1 | MTV041 | AL021958 Mycobacte |

ALIGNMENTS

| | | | | | |
|------------|--|---------|------|-----|-------------|
| RESULT 1 | RNU94330 | 2432 bp | mRNA | ROD | 06-MAY-1997 |
| RNU94330 | Rattus norvegicus osteoprotegerin (OPG) mRNA, complete cds. | | | | |
| LOCUS | U94330 | | | | |
| DEFINITION | g2072180 | | | | |
| ACCESSION | | | | | |
| NID | | | | | |
| KEYWORDS | Norway rat. | | | | |
| SOURCE | Rattus norvegicus | | | | |
| ORGANISM | Eukaryotae; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. | | | | |
| REFERENCE | 1 (bases 1 to 2432) | | | | |
| AUTHORS | Simonet,W.S., Lacey,D.L., Dunstan,C.R., Kelley,M., Chang,M.S., Luthy,R., Nguyen,H.O., Wooden,S., Bennett,L., Boone,T., Shimamoto,G., DeRose,M., Elliott,R., Colombero,A., Tan,H.L., Trail,G., Sullivan,J., Davy,E., Bucay,N., Renshaw-Gegg,L., Hughes,T.W., Hill,D., Pattison,W., Campbell,P., Sander,S., Van,G., Tarpley,J., Derby,P., Lee,R., Amgen EST program and Boyle,W.J. | | | | |
| TITLE | Osteoprotegerin: a novel secreted protein involved in the regulation of bone density | | | | |
| JOURNAL | Cell 89 (2), 309-319 (1997) | | | | |
| MEDLINE | 97262071 | | | | |


```

REFERENCE 2 (bases 1 to 2432)
AUTHORS Boyle,W.J.
TITLE Direct Submission
JOURNAL Submitted (18-MAR-1997) Department of Cell Biology, Angen, Inc.,
1840 Dehavilland Drive, Thousand Oaks, CA 91320, USA
FEATURES
source
1..2432
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/cclone="pMOB B1.1"
124..1329
/gene="OPG"
124..1329
/gene="OPG"
/codon_start=1
/product="osteoprotegerin"
/db_xref="PID:g2072181"
/translation="MNKWLCCALLVFLDIIEWTTOETPPPKYHYDPETGQLLDCDK
APGTYLKQHTVRRKTLVCPDHSYTDWSHTSDEVCYSPVKELQSVKQECNRTHN
RVCEEGRYLEIEFLKRNHSCPPGLGYLAGTPERNVCKRCPDGFSGEITSKAPC
RKHTNCSIGLLIQLQGNATHDNVSGNREATQNGIDVTICEEAFRFAYPTKIIPN
WLSVLVSLPGTKVNAESVERIKRHSQEQTFOLLKWKHQNRPQEMVKKIIQIDIL
CESSVQRHGLSHNLTFEQLALMESLPGKISPEIERKTKRCKPSEOLLKLLSWRI
KNGDDTLKGLMYALKHLKAYHFPKTVTHSLRKTIRFLSHFTMYRLYOKLFLEMIGNQ
VQSVKISCL"
BASE COUNT 682 a 564 c 562 g 624 t
ORIGIN

Query Match 18.1%; Score 83.2; DB 13; Length 2432;
Best Local Similarity 53.4%; Pred. No. 2 1e-08;
Matches 175; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 10 tgcgcgtactgcaacgtctctctggtgggagcgtgaggagagcagggccttgccacgcc 69
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 TGGGTGTACTGACGCCCGCTGTGCAAGAACTGCAGACCGTGAACAGGAGTGCAACCCG 420
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 70 accacaacgctgcctgcctgcgcacacggccttcttcgcgcacgctgttctgtctg 129
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 ACCCAACCGAGTGTGCGAATGTGAGGAAGGCGCTACTCGAGCTCGAATTCGCTTG 480
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 130 gagcagcatcgtgtccacctggtgcggcgtgattgccccgggacccccagcagaca 189
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 AAGCACCAGGAGTGTCCCGCAGGCTGTGGTGTGCTCCAGGCTGGGACCCAGAGCGAAC 540
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 190 acgcaatgcaacgctgccccccagggcacctctcttcagcagcagctccagctcagagcag 249
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 ACGGTTTGCAAAAGATGTCCGGATGGTTCTTCTCAGGTGAGACGTCATCGAAAGCACCC 600
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 250 tgccagccccacgcgaactgcagcgcctggcctggcctgacctcaatgtgcccagctctcc 309
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 TGTAGAAACACCAACCTGACGCTCACCTGGCCCTCCTGCTAATTCAGAAAGAAATGCA 660
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 310 tcccatgacacctgtgacaccagctgca 337
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 661 ACACATGACATGATGTTCCGGAACA 688
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
LOCUS MMU94331 1325 bp mRNA ROD 06-MAY-1997
DEFINITION Mus musculus osteoprotegerin (OPG) mRNA, complete cds.
ACCESSION U94331
NID 92072182
KEYWORDS house mouse.
SOURCE Mus musculus
ORGANISM Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 1325)
AUTHORS Simonet,W.S., Lacey,D.L., Dunstan,C.R., Kelley,M., Chang,M.S.,
Luthy,R., Nguyen,H.Q., Wooden,S., Bennett,L., Boone,T.,

```

```

Shimamoto,G., DeRose,M., Elliott,R., Colombero,A., Tan,H.L.,
Trall,G., Sullivan,J., Davy,E., Bucay,N., Renshaw-Gegg,L.,
Hughes,T.M., Hill,D., Pattison,W., Campbell,P., Sander,S., Van,G.,
Tarpley,J., Derby,P., Lee,R., Angen EST Program and Boyle,W.J.
Osteoprotegerin: a novel secreted protein involved in the
regulation of bone density
Cell 89 (2), 309-319 (1997)
JOURNAL 97262071
MEDLINE 2 (bases 1 to 1325)
REFERENCE Boyle,W.J.
AUTHORS Direct Submission
TITLE Submitted (18-MAR-1997) Department of Cell Biology, Angen, Inc.,
JOURNAL 1840 Dehavilland Drive, Thousand Oaks, CA 91320, USA
FEATURES
Location/Qualifiers
source
1..1325
/organism="Mus musculus"
/db_xref="taxon:10090"
91..1296
/gene="OPG"
91..1296
/gene="OPG"
/codon_start=1
/product="osteoprotegerin"
/db_xref="PID:g2072183"
/translation="MNKWLCCALLVLDIIEWTTOETPPPKYHYDPETGQLLDCDK
APGTYLKQHTVRRKTLVCPDHSYTDWSHTSDEVCYSPVKELQSVKQECNRTHN
RVCEEGRYLEIEFLKRNHSCPPGSGVQAGTPERNVCKRCPDGFSGEITSKAPC
IKHTNCSIGLLIQLQGNATHDNVSGNREATQNGIDVTICEEAFRFAYPTKIIPN
WLSVLVSLPGTKVNAESVERIKRHSQEQTFOLLKWKHQNRPQEMVKKIIQIDIL
CESSVQRHGLSHNLTFEQLALMESLPGKISPEIERKTKRCKPSEOLLKLLSWRI
KNGDDTLKGLMYALKHLKTSHPKTVTHSLRKTIRFLSHFTMYRLYOKLFLEMIGNQ
VQSVKISCL"
BASE COUNT 381 a 329 c 335 g 279 t 1 others
ORIGIN

Query Match 16.5%; Score 75.6; DB 13; Length 1325;
Best Local Similarity 52.2%; Pred. No. 8.7e-07;
Matches 168; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

QY 16 tactcagcgtcctctcggggagcgtgaggagagcagggccttgccagcccccac 75
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 334 TATTCACGCCAGTGTGCAAGAACTGCAGTCCGTTGAAGCAGGAGTGCAACGCCACCCAC 393
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 76 aacgtgctgcctgcgcgcacgcgctctcttcgcgcgcgctggttctgtctgtgagcac 135
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 394 AACCGAGTGTGTGAGTGTGAGGAAGGCGTTACCTGGAGATCGAATTCGCTTGAAGCAC 453
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 136 gcactgctgcaacctggtgcggcgctgattgccccgggacccccagcagaacacgcag 195
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 454 CGGAGCTGTCCCGCGGCTCCGGCGTGGTGAAGCTGGAAACCCAGAGCGAAACACAGTT 513
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 196 tgccagcgcgtgccccccagggcacctcttcagccagcagctccagctcagcagtgccag 255
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 514 TCGAAAAAATGTCCAGATGGGTCTCTCAGGTGAGATTCATCGAAAGACCCCTGTATA 573
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 256 ccccccgcgaactgcagcggccctggcctggccctcaatgtgcccagctcttccctccat 315
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 574 AAACACACGAACTGCAGCACATTTGGCCCTCCTGCTAATTCAGAAAGAAATGCAACACAT 633
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 316 gacacctgtgcaaccagctgca 337
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 634 GACAACGTGTGTTCGGGAACA 655
||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
AB013898 2818 bp mRNA ROD 22-SEP-1998
LOCUS Mus musculus mRNA for osteoclastogenesis inhibitory factor,
DEFINITION complete cds.
ACCESSION AB013898
NID 93142238
KEYWORDS osteoclastogenesis inhibitory factor.

```



```

SOURCE      Mus musculus (strain:NIH Swiss) cell_line:NIH3T3 CDNA to mRNA.
ORGANISM
REFERENCE
AUTHORS     Mizuno.A., Murakami.A., Nakagawa.N., Yasuda.H., Tsuda.E.,
             Morinaga.T. and Higashio.K.
TITLE       Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
             gene and its expression in embryogenesis
JOURNAL     Gene 215 (2), 339-343 (1998)
MEDLINE     98382527
REFERENCE   2 (bases 1 to 2818)
AUTHORS     Mizuno.A. and Yasuda.H.
TITLE       Direct Submission
JOURNAL     Submitted (13-MAY-1998) to the DDBJ/EMBL/GenBank databases. Atsuko
             Mizuno, Snow Brand Milk Products Co., Ltd., Research Institute of
             Life Science; 519 Shimo-Ishibashi, Ishibashi-machi, Tochigi
             329-0512, Japan (E-mail:fvbd7042@mb.infoweb.or.jp,
             Tel:0285-52-1331, Fax:0285-53-1314)
COMMENT     On May 20, 1998 this sequence version replaced gi:3135221.
FEATURES
Source      1..2818
             /organism="Mus musculus"
             /strain="NIH Swiss"
             /db_xref="taxon:10090"
             /cell_line="NIH3T3"
             238..1443
             /note="mOCIF"
             /codon_start=1
             /product="osteoclastogenesis inhibitory factor"
             /db_xref="PID:d1029209"
             /db_xref="PID:g3135222"
             /translation="MNKWLCCALLVLDIETWTOETLPPKVLHYDPETHQLLCDKC
             APTGLKQHTVRRITLPCPDHSDYSDSWHSDSCVYSPVCKELQVYKQBCNRTHN
             RVCEEGRIEIEFLKHSRCPGGVGVQAGTPEPTVCKKCPDGFSGSTSSKAPC
             RKHTDCSTFGLLIQGNATHDNCNREATQKCGIDVTLCEEAFFFAVPTKTPN
             WLSVLDPGLPTKVNAESVERIKRHSQEQTFOLLKWKHQRNQEMVKKIQDIDL
             CESSVORHLGHANLTQOLRMALMESLPGRKISPEIERTRKTKSSEQLKLLSLWRI
             KNGDDTLKGLMYALKHLKTSHPKTVTHSLRKTFRFLHSFTMYRLYQKLFLEMIGNQ
             VQSVKISCL"
BASE COUNT 844 a 603 c 611 g 760 t
ORIGIN
Query Match 16.1%; Score 74; DB 13; Length 2818;
Best Local Similarity 51.9%; Pred.No.1.6e-06;
Matches 167; Conservative 0; Mismatches 155; Indels 0; Gaps 0;
QY 16 tactgcaacgtctctgcggggagcggtgaggaggagcagcggtctgccaagccacccac 75
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 481 TATTGCAGCCACTGTGCAAGGAACGTCAGTCGCGTGAAGCAGAGTGCAACCCGACCCAC 540
QY 76 aacgctgcctgcgcgcgcgcgcacccgctcttcgcgcacgctggttctgcttgagcac 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 541 AACCGAGTGTGTGAGTGTGAGGAAGGCGGTACCTGAGATCGAATCTGCTTGAAGCAC 600
QY 136 gcatcgtgccactgtagtcgcgcgcgattgccccgggacccccagccagacagcag 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 601 CGAGAGTGTCCTCCCGGCTCCGCGGTGTCGAAGCTGGAAACCCAGACCAACACAGTT 660
QY 196 tgcagccgctgccccagcactctctcagccagcagcagctccagcagcagtgccag 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 661 TGCAGAAATGTCAGATGGTTCTTCTCAGGTGAGACITCATCGAAGACCCCTGTAGA 720
QY 256 cccacccgcaactgcagcgcgcctggcctggcctcaatgtgccagctcttctccccc 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 721 AAACACACGAGCTGCAGCACATTTGGCTCTCTGCTAATTTCAGAAAGGAATGCAACACAT 780
QY 316 gacacccctgtgaccagctgca 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 781 GACAACGTGTGTTCCGGAAACA 802

```

```

RESULT      4
LOCUS       AB002146
DEFINITION Homo sapiens mRNA for osteoclastogenesis inhibitory factor (OCIF),
             complete cds.
ACCESSION   AB002146
NID         g3107916
KEYWORDS    Homo sapiens lung fibroblast cell_line:IMR-90 cDNA to mRNA,
             clone_lib:lambda ZAP EXPRESS cDNA library of IMR-90 cells
SOURCE      clone:lambda Olf10.
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
             Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (sites)
AUTHORS     Yasuda.H., Shima.N., Nakagawa.N., Mochizuki.S., Yano.K., Fujise.N.,
             Sato.Y., Goto.M., Yamaguchi.K., Kuriyama.M., Kanno.T., Murakami.A.,
             Tsuda.E., Morinaga.T. and Higashio.K.
TITLE       Identity of osteoclastogenesis inhibitory factor (OCIF) and
             osteoprotegerin (OPG): a mechanism by which OPG/OCIF inhibits
             osteoclastogenesis in vitro
JOURNAL     Endocrinology 139 (3), 1329-1337 (1998)
MEDLINE     98151033
REFERENCE   2 (bases 1 to 1206)
AUTHORS     Yasuda,H.
TITLE       Direct Submission
JOURNAL     Submitted (25-MAR-1997) to the DDBJ/EMBL/GenBank databases.
             Hisataka Yasuda, Snow Brand Milk Products Co., Ltd., Research
             Institute of Life Science; 519 Shimo-Ishibashi, Ishibashi-machi,
             Tochigi 329-05, Japan (E-mail:fvbd7042@mb.infoweb.or.jp,
             Tel:0285-52-1331, Fax:0285-53-1314)
COMMENT     Sequence updated (28-Apr-1998).
FEATURES
Source      1..1206
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /cell_line="IMR-90"
             /cell_type="fibroblast"
             /clone_lib="lambda Olf10"
             /clone_lib="lambda ZAP EXPRESS cDNA library of IMR-90
             cells"
             /tissue_type="lung"
sig_peptide 1..63
             /note="signal peptide (aa -21 to -1)"
             /evidence=experimental
CDS         1..1206
             /codon_start=1
             /product="osteoclastogenesis inhibitory factor (OCIF)"
             /db_xref="PID:d1026850"
             /db_xref="PID:g3107917"
             /translation="MNKLLCCALVFLDISIKWTTQETPPPKYLHYDEETSHQLCDKC
             PPGTYLKQHTAKWKTVCAPCDPDHYITDSWHTSDCLYCSVPVCKELQVYKQBCNRTHN
             RVCEKGEGRILEIEFLKHSRCPGGVGVQAGTPEPTVCKKCPDGFSGSTSSKAPC
             RKHTNCVFGLLITOKGNATHDNCNREATQKCGIDVTLCEEAFFFAVPTKTPN
             WLSVLVDNLPGLPTKVNAESVERIKRHSQEQTFOLLKWKHQRNQDQDVLVKKIQDIDL
             CENSQVRIHGHNLTFFEOLRSIMESLPGRKISPEIERTRKTKSSEQLKLLSLWRI
             KNGDDTLKGLMHALKHSKTHFPKTVTQSLUKKTIIRFLHSFTMYRLYQKLFLEMIGNQ
             VQSVKISCL"
             9
variation    9
             /replace="c"
variation    22
             /replace="t"
mat_peptide 64..1203
             /note="mature OCIF (aa 1 to 380)"
             /evidence=experimental
             /product="osteoclastogenesis inhibitory factor (OCIF)"
misc_feature 67..189
             /note="cysteine-rich domain 1"
misc_feature 190..318
             /note="cysteine-rich domain 2"
misc_feature 319..429
             /note="cysteine-rich domain 3"
misc_feature 430..558

```



```
QY 10 tgcgcctactgcaacgtccctctcggggagcggtgagggagggcagcggttgccagcc 69
Db 332 TGTCTACTGACGCCCGTGTCAAGGAGCTGCAGTACGTACGTCAAGCAGGAGTGCATCGC 391
QY 70 acccaaacgcgtgctgcgcgtgcgcgcacgcggttcttcgcgcacgctgtttctctgtg 129
Db 392 ACCCACAAACGCGTGTGCGAATGCAAGGAAGGGCGCTACCTTGAGTAGAGTGTCTGCTG 451
QY 130 gacacgcacgtgctcacctggtgcgcggcggtgattgcccccgggcagccccagccagaac 189
Db 452 AAACATAGGAGCTGCCCTCTCGATTGGATTTGGATGGTGAAGCTGGAACCCAGAGCGAAAT 511
QY 190 acgcagtgccagcgcgtccccccagggcacctcttcagccagcagctccagctcagagcag 249
Db 512 ACAGTTTGCAAAAGATGTCAGATGGGTCTTCTCAATGAGAGCTCATCTAAGGACCC 571
QY 250 tgcagccccacgcgcgaactgcgcgcgcctggcctgcctcaatgtgcagcgctttcc 309
Db 572 TGTAGAAAACACACAAATTGCAGTGTCTTTGGTCTCTCTGCTAACTCAGAAAGGAATGCA 631
QY 310 tccatgacaccctgtgcacacagctgactg 340
Db 632 ACACAGCAACATATGTTCCGGAACAGTG 662

RESULT 8
MMTNR2A MMTNR2A 1388 bp RNA ROD 17-FEB-1997
LOCUS M.musculus tumor necrosis factor receptor 2 mRNA.
DEFINITION X76401
ACCESSION 9433830
NID 9433830
KEYWORDS tumour necrosis factor receptor.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 1388)
AUTHORS Powell,E.E.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-1993) E.E. Powell, c/o John Todd, Level 6
Nuffield Dept of Surgery, The John Radcliffe Hospital, Oxford, UK
REFERENCE 2 (bases 1 to 1388)
AUTHORS Powell,E.E., Wicker,L.S., Peterson,L.B. and Todd,J.A.
TITLE Amino acid variation in the tumor Necrosis factor receptor 2 is
linked to autoimmune diabetes in NOD mice
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1388)
AUTHORS Powell,E.E., Wicker,L.S., Peterson,L.B. and Todd,J.A.
TITLE Allelic variation of the type 2 tumor necrosis factor receptor gene
Mamm. Genome 5 (11), 726-727 (1994)
MEDLINE 95178848
FEATURES
source Location/Qualifiers
1..1388
/organism="Mus musculus"
/strain="NOD"
/db_xref="taxon:10090"
/chromosome="4 (distal region)"
1..1380
/gene="murine tumour necrosis factor receptor 2"
<1..1380
/gene="murine tumour necrosis factor receptor 2"
/codon_start=1
/db_xref="PID:g433831"
/translation="LWATGHTVPAQVLPYKPEGYECOISQEYYDRKAQMCCAKCP
PGQYVKNFKNTSDTYCDAEASMYQVWQNFRTCLSSSGSTDDVETRACTKQNR
VCACEGRYCALKTHSGRCQMRLSKCGPFGVASSRAPGNVLKACAPTFSDTT
STDVCRPHRICISILAPGNASTDVAPESTPLSAIPRTLYVSOEPTFSPLODEP
GPSQPSILTSLSGTPILIQSTKGGISLPIGLIVGTSIGLMLGVNCFILVQRKK
PCLORDAKVPHVPDKSDAVGLEQHLHTAPSSSSSSLESSASAGDRRAPPGGHP
QARVMAEGSGEARSSISDSHSGHGHVNTVIVNCSHSSSHSSQCSQASATV
GPDFAKPSAPKDEQYFPFQEQECPSQSPYETTETLQSEKPLPLGLVPMGMKPSQAGW
FDQIAVKVA"
```

```
variation
260
/gene="murine tumour necrosis factor receptor 2"
/notes="Ser to Thr"
/replace="c"
278
/gene="murine tumour necrosis factor receptor 2"
/notes="Thr to Ile"
/replace="t"
489
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="t"
802
/gene="murine tumour necrosis factor receptor 2"
/notes="Phe to Ile"
/replace="a"
921
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="c"
975
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="c"
1034
/gene="murine tumour necrosis factor receptor 2"
/notes="Ser to Phe"
/replace="t"
1047
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="t"
1143
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="t"
1173
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="a"
1262
/gene="murine tumour necrosis factor receptor 2"
/notes="Thr to Cys"
/replace="g"
1317
/gene="murine tumour necrosis factor receptor 2"
/notes="silent"
/replace="g"

BASE COUNT 327 a 417 c 371 g 273 t
ORIGIN
```

```
Query Match 11.5%; Score 52.8; DB 13; Length 1388;
Best Local Similarity 52.2%; Pred. No. 0.043;
Matches 117; Conservative 0; Mismatches 107; Indels 0; Gaps 0;
```

```
QY 124 tgcctggagcagcatggtgtccacctgtgctgcggcggtgattgccccgggcacccccagc 183
Db 370 TGCATGAGGCTGAGCAAGTGC GGCCCTGCTTGGAGTGGCCAGTTCAGAGGCCCAAAAT 429
QY 184 cagaacacgcagtgccagcgcgtgccccccagggcaccttctcagccagcagctccagctca 243
Db 430 GGAATGTGCTATGCAAGGCTGTGCCCCAGGACGTTCTCTGACACCATCATCCACA 489
QY 244 gacagtgccagccacccacgcgaactgcacggccctggcgccctcaatgtgcccaggc 303
Db 490 GATGTGTGACAGGCCACCCGCCATCTGTAGCATCTCTGGCTATTCCCGGAAATGCAAGCACA 549
QY 304 tctctcccatgacacctgtgcacagctgcactggtctccc 347
Db 550 GATGCAGTCTGTGCGCCGAGTCCCAACTCTTAAGTGCATCCC 593
```

```

RESULT 9
MUSTNFR1 3796 bp mRNA ROD 01-AUG-1991
LOCUS Murine tumor necrosis factor I receptor (TNFR-1) mRNA, complete cds
DEFINITION
ACCESSION M59378
NID g202094
KEYWORDS tumor necrosis factor receptor.
SOURCE Mus musculus lymphoid cDNA to mRNA.
ORGANISM Mus musculus
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 3796)
AUTHORS Goodwin,R.G., Anderson,D.M., Jerzy,R., Davis,T., Brannan,C.I.,
Copoland,N.G., Jenkins,N.A. and Smith,C.A.
TITLE Molecular cloning and expression of the type 1 and type 2 murine
receptors for tumor necrosis factor
JOURNAL Mol. Cell. Biol. 11, 3020-3026 (1991)
MEDLINE 91246168
FEATURES
Location/Qualifiers
source 1..3796
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_line="7B9"
/cell_type="helper T-cell"
/tissue_type="lymphoid"
/map="Chromosome 4"
43..108
/gene="TNFR-1"
43..1467
/gene="TNFR-1"
/product="tumor necrosis factor receptor 1"
/db_xref="PID:g202095"
/translation="MAPAALWALVLFELQWLWATGHTVPAQVILTPYKPEGYECQISQ
EYDRKAQMCACKPPQYVYKFCNKTSDTVACDCEASMTQVWNOFRTCLSGSSSCT
TDQVEIRACTKQNRVCACAGYKCALKTSHGSCROCRLSKGPGFGVASSRAPNGN
VLCKACAPGTFSDTTSIDVCRPHRICSLAIPGNASTDAVCAPESTLSAIPRTLTV
SQPEPTRSQBLDQEPGSPQTSILTSIGSTPIIEQSTKGISLPIGLIVGTVSLGLLM
LGLVNCILVQRKKKPCLODKAVPHVPDEKSDQAVGLEQQHLLITAPSSSSSLES
SASAGRRAPPGHPQARVMAEQGFQEARSSRISDSHSGSHGTHVNTVCIVNVCS
SDHSSQSSQASATVGDPAKPSAPKDEQVPFSQECPSQSPCTTETLQSHKPLP
LGVPMGMKPSQAGWFDQIAKVA"
43..1467
/gene="TNFR-1"
mat_peptide 109..1464
/gene="TNFR-1"
/product="tumor necrosis factor receptor 1"
BASE COUNT 863 a 1052 c 981 g 900 t
ORIGIN
Chromosome 4.
Query Match 11.2%; Score 51.2; DB 13; Length 3796;
Best Local Similarity 51.8%; Pred. No. 0.075;
Matches 116; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
Qy 124 tgcctggagcagcgcgcgtgtccacctggtgcccggcggtgattgccccgggcacccccagc 183
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 457 TGCATGAGGCTGAGCAAGTGGCGCCCTGGCTTCGGAGTGGCCAGTTCAAGAGCCCAAAAT 516
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 184 cagaacacgcagtcgcagccgtgccccaggcaccttctcagcagcagctccagctca 243
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 517 GGAATGTGCTATGCAAGGCTGTGCGCCAGGCGGTCTCTGACACCAACATCATCCACT 576
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 244 gagcagtgccagcccccacacgacccctggcctggcctggcctcaatgtgccaggc 303
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 577 GATGTGTGAGCCGCCACCGCATCTGTAGCATCTGGTATTCGCGGAATGCAAGCACA 636
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 304 tcttctcccatgacacctgtgcacacgctgcacctgacctgacctcc 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 637 GATGCAGTCTGTGGCGCGGAGTCCCAACTCTAAGTGCCATCCC 680
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

BASE COUNT 347 a 459 c 408 g 291 t
ORIGIN
Query Match 11.2%; Score 51.2; DB 13; Length 1505;
Best Local Similarity 51.8%; Pred. No. 0.09;
Matches 116; Conservative 0; Mismatches 108; Indels 0; Gaps 0;
Qy 124 tgcctggagcagcgcgcgtgtccacctggtgcccggcggtgattgccccgggcacccccagc 183
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 455 TGCATGAGGCTGAGCAAGTGGCGCCCTGGCTTCGGAGTGGCCAGTTCAAGAGCCCAAAAT 514
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 184 cagaacacgcagtcgcagccgtgccccaggcaccttctcagcagcagctccagctca 243
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 515 GGAATGTGCTATGCAAGGCTGTGCGCCAGGCGGTCTCTGACACCAACATCATCCACT 574
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 244 gagcagtgccagcccccacacgacccctggcctggcctggcctcaatgtgccaggc 303
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 575 GATGTGTGAGGCGCCACCGCATCTGTAGCATCTGTGCTATTCGCGGAATGCAAGCACA 634
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 304 tcttctcccatgacacctgtgcacacgctgcacctgacctgacctcc 347
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 635 GATGCAGTCTGTGGCGCGGAGTCCCAACTCTAAGTGCCATCCC 678
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
AF019046
LOCUS AF019046 2116 bp mRNA ROD 22-NOV-1997
DEFINITION Mus musculus receptor activator of nuclear factor kappa B (RANK)
mRNA, complete cds.
ACCESSION AF019046
NID g2612919
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus

```



```
RESULT 13
AF068868      1968 bp      mRNA      PRI      07-SEP-1998
LOCUS
DEFINITION    Homo sapiens TNFR-related death receptor-6 (DR6) mRNA, complete
               cds
ACCESSION     AF068868
NID           93549262
KEYWORDS
SOURCE        human.
ORGANISM      Homo sapiens
               Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
               Primates; Catarrhini; Homnidae; Homo.
REFERENCE     1 (bases 1 to 1968)
AUTHORS       Pan,G., Ni,J., Yu,G. and Dixit,V.M.
TITLE         Vincenz,C., Aggarwal,B.B. and Dixit,V.M.
               Identification and functional characterization of DR6, a novel
               death domain-containing TNF receptor
JOURNAL       Unpublished
REFERENCE     2 (bases 1 to 1968)
AUTHORS       Pan,G., Ni,J., Yu,G. and Dixit,V.M.
TITLE         Direct Submission
JOURNAL       Submitted (28-MAY-1998) Pathology, University of Michigan Medical
               School, 1301 Catherine Road,, Ann Arbor, MI 48109, USA
FEATURES
source        1..1968
               Location/Qualifiers
               /organism="Homo sapiens"
               /db_xref="taxon:9606"
gene          1..1968
               /gene="DR6"
CDS           1..1968
               /gene="DR6"
               /codon_start=1
               /product="TNFR-related death receptor-6"
               /db_xref="PID:g3549263"
               /translation="MGTSFSSSTALASCSRIARRATAMAGSLLLGLFLSTTTAQPE
               QKASNLGTYRHVRATGQVLCDCPKAGTYVSEHCTNITSLRSCSPVGTFRRENG
               IEKHDCQPCWPMIEKLPKCAALTDRETCPPGMFQSNATCAPHTVCPVGVGRKKG
               TETEDVRGQKARGTFSDVPKMKCAATYDCLSONLVVVKPTETDNVCTLPSPFS
               SSTSPGTAIPRPEHMETHEVPSSTVVPKGMNSTENSSASVRPKVLSLSIQECTVP
               DNTSSARKEDVNTKLPNLOVYVHOOGPHRHILKLLPSMEATGEEKSSTPIKPKRG
               HPRQLNHFEDLNHLPMWVLFLLLVVYVCSIRSKSLTKKGRDPPSAIVEKA
               GLKSMFTYQNEKWIYCYNGHGDIDLKLVAAQGVQSKWDIYQFLCNASEREVAAFSN
               GYTDHSPALALQHWITRGPEASLAQLISALRHRNDVVEKIRGLMEDTQLTQLETDK
               LALPMSPSLSPSPSPSPNAKLENSALLTVEPQDKNKGFFVDESEPLLRCDSTSSG
               SSALRSGSFITKEKNDVLRVRLDPCDLOQIFDDMLHFLNPEELRVIEEIPQAEKD
               LDRLEILGVKSQEAQSLLDLSVYSHLPDL"
BASE COUNT   479 a 583 c 512 g 394 t
ORIGIN
Query Match      10.3%; Score 47.2; DB 42; Length 1968;
Best Local Similarity 45.8%; Pred. No. 0.57;
Matches 163; Conservative 0; Mismatches 193; Indels 0; Gaps 0;
QY 2 tggagcgctgcgcgtactgcaacgctctctgcggggagcggtgagggagggcagcggtt 61
Db 308 TAGAGAAATGCCATGACTGTAGTCAGCCATGCCCATGGCCATGATGAGAAATTACCTT 367
QY 62 gccacgcaccacaacacgctgctgcgctgcgcgacccggtctcttcgcgcagcgtggtt 121
Db 368 GTGCTGCTTGACTGACCGAGAAATGCACTTGCCACACCTGGCATGTTCACGCTTAAGCCTA 427
QY 122 tctgcttgagacgacgcatcgctccacctggtgctcgccgctgattgccccggcaccacca 191
Db 428 CTTGTGCCCCCATACAGGTGTGTCCTGTGGGTGGGGGTGTGCGGGAAGAAGGACAGAGA 487
QY 182 gccacacacgcagtgccagcgctgccccccagccacctcttcagccagcagcgtccagct 241
Db 488 CTGAGGATGTGCGGTGTAAAGCAGTGTGCTCGGGGTACCTTCTCAGATGTGCCCTTCTAGTC 547
QY 242 cagagcagtgccagccccccaccccaactgcacggccctggccctgcccccaatgtgccag 301
Db 548 TGATGAATGCAAGCATACAGACTGTCTGAGTCAGAACCTGGTGGTGGTGGTGGTGGTGGTGG 607
```

```
QY 302 gctcttcctcccatgacaccctgtgcaccagctgcactggcttccccctccagcacc 357
Db 608 GGACCAAGAGACAGACAGACAGCTGTGTGGACACTCCCGTCTCTCCAGCTCCACC 663

RESULT 14
A26415      2339 bp      DNA      PAT      26-APR-1995
LOCUS
DEFINITION    CDNA fragment for (75kd TNF-BP) tumor necrosis factor binding
               protein from patent EP0417563.
ACCESSION     A26415
NID           g904970
KEYWORDS       unidentified.
SOURCE        unidentified.
ORGANISM      unclassified.
REFERENCE     1 (bases 1 to 2339)
AUTHORS       Brockhaus,M., Dembic,Z., Gentz,R., Lesslauer,W., Loetscher,H. and
               Schlaeger,E.J.
TITLE         TNF-binding proteins
JOURNAL       Patent: EP 0417563-A 27 20-MAR-1991;
               F. HOFFMANN-LA ROCHE AG
FEATURES
source        1..2339
               Location/Qualifiers
               /organism="Artificial sequences"
               <1..1179
               /note="pid:el52230"
               /codon_start=1
               /product="75kd TNF-BP"
               /db_xref="PID:g904971"
               /translation="SDSYCDSCEDSTYTQLNNWVPECLSCGRSCSSDOVETOACTREQ
               NRICTRGWCALSKQSGRCRLCAPLPCRFGFVGARGTETSDVPCARGTSNT
               TSSDIPRHOICNVVAIPGNASRDVACTSTSPTRSMAPVHLPQVPVRSQHTOPS
               PEPSTPSTSLPMPGPPPAEGSTGDFALPVLIVGTALGLLIGVNVCMPTQVK
               KQPLQREAKVPHLPADKQTGPEQHLLITAPSSSSLESSSALSADRRAPTRN
               OQAPGVASGAGEARASTGGSADSPPGHGTQVNVTCIVNVSSSDHSSOCSSOASS
               TMGDTDSSESPEKDEQVFPFSKEECARSQLPTEPPLLSTEEKPLPLGVDPAGMKPS
               "
BASE COUNT     494 a 720 c 685 g 440 t
ORIGIN
Query Match      9.8%; Score 45; DB 6; Length 2339;
Best Local Similarity 49.0%; Pred. No. 1.6;
Matches 120; Conservative 0; Mismatches 125; Indels 0; Gaps 0;
QY 51 ggcacgggttgccacgccaccacacacgctgctcgctgcgcgacgggtcttcgc 110
Db 126 GGAACAGAACCGCATCTGCACCTGCAGGCCCGCGGTGCTGCTGCGCTGAGCAAGCAGGA 185
QY 111 gcacgtgtttctgcttggagcagcagctgtgtccacctggtgcggtggtgattgcccc 170
Db 186 GGGGTGCGGCTGTGCGCGCGCTGCCGAAGTGCCTGCGCGGGCTTCGGCTGCGCAGACC 245
QY 171 gggacccccccagcagaacacgacgctgcgcgctgccccccagggcaccttccagccag 230
Db 246 AGGAACGTAAACATCAGACGTGGTGTGCAAGCCCTGTGCCCCGGGACGTTCTCCAACAC 305
QY 231 cagctcagctcagagcagtgccagccccccacccagcaactcagcgccctggcctggccct 290
Db 306 GACTTCATCCAGGATATTGTGAGGCCCCACCACGATCTCTAAGTGGTGGCCATCCCTGG 365
QY 291 caatg 295
Db 366 GAATG 370

RESULT 15
A43530      1600 bp      DNA      PAT      06-MAR-1997
LOCUS
DEFINITION    Sequence 1 from Patent WO9506723.
ACCESSION     A43530
```


FT protein 22..399 /note= "Mature OCIF-CL"
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB; T33172.
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 62; Page 117-119; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CL in which amino acids 379-380 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 399 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 399;
Best Local Similarity 41.5%; Pred. No. 8.8e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYNVLCGEREEERACHATHNRACRGTGFFAHAGFCLEHASCPGAGVIAPGTGS 61
Db 78 DECLYSPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHKRSCPPGFGVVOAGTPE 137
QY 62 ONTQCPCPPGTFSASSSSSEOCOPHRNCTALGLALNVPGSSSHDTLCTCTGFPPLSTRV 121
Db 138 RNVCKRCPPDGFSTNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNI---CSGNSESTQK 194
QY 122 PGAE--CERAVIDF 134
Db 195 CGIDVTLCEEAFFR 209

RESULT 8
R99943
ID R99943 standard; Protein; 351 AA.
AC R99943;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CC.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT protein 22..351
FT /note= "Mature OCIF-CC"
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB; T33173.
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 65; Page 119-121; 183pp; Japanese.

CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CC in which amino acids 331-380 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 351 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 351;
Best Local Similarity 41.5%; Pred. No. 7.6e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYNVLCGEREEERACHATHNRACRGTGFFAHAGFCLEHASCPGAGVIAPGTGS 61
Db 78 DECLYSPVKELQYVQECNRTNHRVCECKEGRYLEIEFCLKHKRSCPPGFGVVOAGTPE 137
QY 62 ONTQCPCPPGTFSASSSSSEOCOPHRNCTALGLALNVPGSSSHDTLCTCTGFPPLSTRV 121
Db 138 RNVCKRCPPDGFSTNETSSKAPCRKHTNCSVFGLLLTQKGNATHDNI---CSGNSESTQK 194
QY 122 PGAE--CERAVIDF 134
Db 195 CGIDVTLCEEAFFR 209

RESULT 9
R99944
ID R99944 standard; Protein; 272 AA.
AC R99944;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CDD2.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT Peptide 1..21
FT /note= "Signal peptide"
FT protein 22..272
FT /note= "Mature OCIF-CDD2"
PN WO9626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB; T33174.
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 68; Page 121-122; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CDD2 in which amino acids 252-380 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 272 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 272;

Best Local Similarity 41.5%; Pred. No. 5.8e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYCNVLCGEREERACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DECLYCSVPCKELQVQKCEKRNTHRVCEKRGYLEIEFCLKHRSCTPPGFGVQAGTPE 137

QY 62 QNTQCPCPPGTFSSASSSSQCPHNRCTALGLALNVPGSSSHDTLCTSGTGFPLSTRV 121
Db 138 RNTVCKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNI---CSGNSESTQK 194

QY 122 PGABE--CERAVIDF 134
Db 195 CGIDVTLCCEAFFRF 209

RESULT 10
R99948
ID R99948 standard; Protein; 393 AA.
AC R99948;
DE 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CBst
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key
FT peptide
FT /note= "Signal peptide"
FT protein
FT /note= "Mature OCIF-CBst"
FT misc_difference 392
FT /label= Gln371Leu
PN W09626217-A1.
PD 29-AUG-1996;
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
PI WPI; 96-402320/40.
DR N-PSDB; T33178.
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 80; Page 126-128; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC Leu and amino acids 373-380 of the mature OCIF protein are deleted.
CC These changes are caused by the introduction of a restriction site in
CC the DNA encoding this protein. The OCIF of the invention has a
CC molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 393 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 393;
Best Local Similarity 41.5%; Pred. No. 8.7e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYCNVLCGEREERACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DECLYCSVPCKELQVQKCEKRNTHRVCEKRGYLEIEFCLKHRSCTPPGFGVQAGTPE 137

QY 62 QNTQCPCPPGTFSSASSSSQCPHNRCTALGLALNVPGSSSHDTLCTSGTGFPLSTRV 121
Db 138 RNTVCKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNI---CSGNSESTQK 194

QY 122 PGABE--CERAVIDF 134
Db 195 CGIDVTLCCEAFFRF 209

RESULT 11
R99933
ID R99933 standard; Protein; 401 AA.
AC R99933;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C21S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key
FT peptide
FT /note= "Signal peptide"
FT protein
FT /note= "Mature OCIF-C21S"
FT misc_difference 277
FT /label= C21S
PN W09626217-A1.
PD 29-AUG-1996;
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
PI WPI; 96-402320/40.
DR N-PSDB; T33163.
DR DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 35; Page 98-100; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C21S in which the 21st Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 401;
Best Local Similarity 41.5%; Pred. No. 8.9e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

QY 2 ERRCYCNVLCGEREERACHATHNRACRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db 78 DECLYCSVPCKELQVQKCEKRNTHRVCEKRGYLEIEFCLKHRSCTPPGFGVQAGTPE 137

QY 62 QNTQCPCPPGTFSSASSSSQCPHNRCTALGLALNVPGSSSHDTLCTSGTGFPLSTRV 121
Db 138 RNTVCKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQKGNATHDNI---CSGNSESTQK 194

QY 122 PGABE--CERAVIDF 134
Db 195 CGIDVTLCCEAFFRF 209

RESULT 12
R99934
ID R99934 standard; Protein; 401 AA.
AC R99934;
DE 22-APR-1997 (first entry)
DE Mutated OCIF, OCIF-C22S.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.

```

OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1..21
FT protein /note= "Signal peptide"
FT protein 22..401
FT misc_difference 277
FT /note= "Mature OCIF-C22S"
FT /label= C22S
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33164.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 38; Page 100-102; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C22S in which the 22nd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 401;
Best Local Similarity 41.5%; Pred. No. 8.9e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

OY 2 ERRCYCNVLCGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
78 DELCYSPVKELQYVKQECNTRNHRVCEKRGYLEIEFCLKHRSCPPGFGVQAGTPE 137

OY 62 QNTQCQPCPPGTFSSASSSSQCPHRNCTALGLALNVPGSSSHDTLCTSCGTGFLPLSTRV 121
Db : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
138 RNTVCKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSSESTQK 194

OY 122 PGAE--CERAVIDF 134
Db : | : ||| : |
195 CGIDVTLCCEAERFR 209

RESULT 13
R99949
ID R99949 standard; Protein; 321 AA.
AC R99949;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-CSph.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1..21
FT protein /note= "Signal peptide"
FT protein 22..321
FT /note= "Mature OCIF-CSph"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33164.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 38; Page 100-102; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-C22S in which the 22nd Cys residue in the
CC mature OCIF protein is substituted by Ser. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
CC cation-exchangers or heparin and its activity is lowered after 10 mins
CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
CC deg.C. OCIF is useful in the control of bone resorption and therefore
CC in the treatment and prevention of disorders of bone resorption, e.g.
CC osteoporosis.
SQ Sequence 401 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 401;
Best Local Similarity 41.5%; Pred. No. 8.9e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

OY 2 ERRCYCNVLCGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
78 DELCYSPVKELQYVKQECNTRNHRVCEKRGYLEIEFCLKHRSCPPGFGVQAGTPE 137

OY 62 QNTQCQPCPPGTFSSASSSSQCPHRNCTALGLALNVPGSSSHDTLCTSCGTGFLPLSTRV 121
Db : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
138 RNTVCKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSSESTQK 194

OY 122 PGAE--CERAVIDF 134
Db : | : ||| : |
195 CGIDVTLCCEAERFR 209

RESULT 14
R99936
ID R99936 standard; Protein; 360 AA.
AC R99936;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCRI.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1..21
FT protein /note= "Signal peptide"
FT protein 22..360
FT /note= "Mature OCIF-DCRI"
FT misc_difference 22..23
FT /note= "Position of deletion, delta 2-42"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33166.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 44; Page 105-107; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-DCRI in which amino acids 2-42 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto

```

```

PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33179.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 83; Page 128-129; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-CSph in which amino acids 298-380 of the mature
CC OCIF protein are replaced by Ser-leu-Asp. These changes are caused by
CC the introduction of a restriction site in the DNA encoding this protein.
CC The OCIF of the invention has a molecular weight by SDS-PAGE of 60 kD
CC under reducing conditions and 120 kD under non-reducing conditions. The
CC protein is adsorbed onto cation-exchangers or heparin and its activity is
CC lowered after 10 mins at 70 deg.C or 30 mins at 56 deg.C, and is lost
CC after 10 mins at 90 deg.C. OCIF is useful in the control of bone
CC resorption and therefore in the treatment and prevention of disorders
CC of bone resorption, e.g. osteoporosis.
SQ Sequence 321 AA;

Query Match 36.8%; Score 309.5; DB 1; Length 321;
Best Local Similarity 41.5%; Pred. No. 6.9e-20;
Matches 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2;

OY 2 ERRCYCNVLCGEREEARACHATHNRACRRTGFFAHAGFCLEHASCPPGAGVIAPGTPS 61
Db : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
78 DELCYSPVKELQYVKQECNTRNHRVCEKRGYLEIEFCLKHRSCPPGFGVQAGTPE 137

OY 62 QNTQCQPCPPGTFSSASSSSQCPHRNCTALGLALNVPGSSSHDTLCTSCGTGFLPLSTRV 121
Db : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
138 RNTVCKRCPDGFSSNETSKAPCRKHTNCSVFGLLLTQGNATHDNI---CSGNSSESTQK 194

OY 122 PGAE--CERAVIDF 134
Db : | : ||| : |
195 CGIDVTLCCEAERFR 209

RESULT 14
R99936
ID R99936 standard; Protein; 360 AA.
AC R99936;
DT 23-APR-1997 (first entry)
DE Mutated OCIF, OCIF-DCRI.
KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
KW osteoporosis.
OS Synthetic.
FH Key Location/Qualifiers
FT peptide 1..21
FT protein /note= "Signal peptide"
FT protein 22..360
FT /note= "Mature OCIF-DCRI"
FT misc_difference 22..23
FT /note= "Position of deletion, delta 2-42"
PN W09626217-A1.
PD 29-AUG-1996.
PF 20-FEB-1996; J00374.
PR 20-FEB-1995; JP-054977.
PR 21-JUL-1995; JP-207508.
PA (SNOW ) SNOW BRAND MILK PROD CO LTD.
PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
DR WPI: 96-402320/40.
DR N-PSDB: T33166.
PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
PT for bone resorption control, esp. treatment of osteoporosis
PS Claim 44; Page 105-107; 183pp; Japanese.
CC This sequence represents a mutated version of the full length
CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
CC sequence represents OCIF-DCRI in which amino acids 2-42 of the
CC mature OCIF protein are deleted. The OCIF of the invention
CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
CC and 120 kD under non-reducing conditions. The protein is adsorbed onto

```


This Page Blank (uspto)

| Result No. | Score | Query % | | | DB | ID | Description |
|------------|-------|---------|--------|--------|--------|--------------------|-------------|
| | | Match | Length | Length | | | |
| 1 | 841 | 100.0 | 300 | 1 | W63622 | Human tumour necro | |
| 2 | 841 | 100.0 | 300 | 1 | W66102 | Amino acid sequenc | |
| 3 | 337 | 40.1 | 170 | 1 | W63623 | Human tumour necro | |
| 4 | 324.5 | 38.6 | 326 | 1 | R99940 | Mutated OC1F, OC1F | |
| 5 | 311.5 | 37.0 | 401 | 1 | R99931 | Mutated OC1F, OC1F | |
| 6 | 309.5 | 36.8 | 327 | 1 | R99941 | Mutated OC1F, OC1F | |
| 7 | 309.5 | 36.8 | 339 | 1 | R99942 | Mutated OC1F, OC1F | |
| 8 | 309.5 | 36.8 | 351 | 1 | R99943 | Mutated OC1F, OC1F | |
| 9 | 309.5 | 36.8 | 351 | 1 | R99943 | Mutated OC1F, OC1F | |
| 10 | 309.5 | 36.8 | 272 | 1 | R99944 | Mutated OC1F, OC1F | |
| 11 | 309.5 | 36.8 | 393 | 1 | R99948 | Mutated OC1F, OC1F | |
| 12 | 309.5 | 36.8 | 401 | 1 | R99933 | Mutated OC1F, OC1F | |
| 13 | 309.5 | 36.8 | 401 | 1 | R99934 | Mutated OC1F, OC1F | |
| 14 | 309.5 | 36.8 | 321 | 1 | R99949 | Mutated OC1F, OC1F | |
| 15 | 309.5 | 36.8 | 360 | 1 | R99936 | Mutated OC1F, OC1F | |
| 16 | 309.5 | 36.8 | 380 | 1 | R99924 | Mature osteoclasto | |
| 17 | 309.5 | 36.8 | 401 | 1 | R99925 | Full length osteoc | |
| 18 | 309.5 | 36.8 | 390 | 1 | R99357 | Human tumour necro | |
| 19 | 309.5 | 36.8 | 401 | 1 | W38345 | Human osteoprotege | |
| 20 | 309.5 | 36.8 | 391 | 1 | W53238 | Human OC1F genome | |
| 21 | 309.5 | 36.8 | 401 | 1 | W53239 | Human OC1F genome | |
| 22 | 306.5 | 36.4 | 401 | 1 | W57636 | Modified TR1 recep | |
| 23 | 305.5 | 36.3 | 401 | 1 | W57635 | Rat osteoprotegeri | |
| 24 | 303.5 | 36.1 | 401 | 1 | W38344 | TR1 receptor prote | |
| 25 | 302.5 | 36.0 | 401 | 1 | R99932 | Mouse osteoprotege | |
| 26 | 302 | 35.9 | 197 | 1 | R99945 | Mutated OC1F, OC1F | |
| 27 | 302 | 35.9 | 187 | 1 | R99950 | Mutated OC1F, OC1F | |
| 28 | 300.5 | 35.7 | 401 | 1 | R99950 | Mutated OC1F, OC1F | |
| 29 | 243 | 28.9 | 461 | 1 | R11001 | 40kD TNF inhibitor | |
| 30 | 243 | 28.9 | 392 | 1 | R11605 | Human 75kD TNF-bi | |
| 31 | 243 | 28.9 | 235 | 1 | W52270 | Tumour necrosis fa | |
| 32 | 243 | 28.9 | 235 | 1 | W59665 | Human soluble tumo | |
| 33 | 241 | 28.7 | 461 | 1 | R11141 | Human TNF-R deduce | |
| 34 | 241 | 28.7 | 485 | 1 | R24016 | Fusion protein TNF | |
| 35 | 241 | 28.7 | 461 | 1 | R24058 | Fibroblast derived | |
| 36 | 241 | 28.7 | 461 | 1 | R72504 | p75 Tumour Necrosi | |
| 37 | 241 | 28.7 | 359 | 1 | R99937 | Mutated OC1F, OC1F | |
| 38 | 237 | 28.2 | 518 | 1 | R51003 | Sequence of a reco | |
| 39 | 232 | 27.6 | 183 | 1 | R77421 | BamTP delta53 nerv | |
| 40 | 218.5 | 26.0 | 474 | 1 | R11142 | TNF-R deduced from | |
| 41 | 215 | 25.6 | 461 | 1 | R51002 | Sequence of human | |
| 42 | 205.5 | 24.4 | 359 | 1 | R99939 | Mutated OC1F, OC1F | |
| 43 | 193 | 22.9 | 143 | 1 | R99946 | Mutated OC1F, OC1F | |

DT 02-DEC-1998 (first entry)

DE Amino acid sequence of tumour necrosis related receptor (TR4).
 KW Human: tumour necrosis related receptor; TR4; agonist; antagonist;
 KW inhibition: chronic; acute; inflammation; arthritis; septicaemia;
 KW autoimmune disease; transplant rejection; stroke; cancer;
 KW Alzheimer's disease.
 OS Homo sapiens.
 PN EP-861850-A1.
 PD 02-SEP-1998.
 PR 20-JAN-1998; 300382.
 PR 04-FEB-1997; US-794796.
 PA (SMIK) SMITHKLINE BEECHAM CORP.
 PI Emery J, Tan KB, Truneh A, Young PR;
 DR WPI: 98-508248/44.
 DR N-PSDB: V07654.
 PT New DNA encoding tumour necrosis related receptor - used to treat
 PT and prevent e.g. inflammation, arthritis, septicaemia, autoimmune
 PT diseases, transplant rejection, infection, stroke, ischaemia, ARDS,
 PT restenosis, AIDS, bone disorders and cancer
 PS Claim 1; Fig 1; 21pp; English.
 CC This is the amino acid sequence of the human tumour necrosis related
 CC receptor (TR4), used in the method of the invention. The TR4 protein
 CC or its agonist can be used to treat a subject in need of enhanced
 CC TR4 polypeptide activity. The antagonist is used to inhibit TR4
 CC polypeptide activity. The active agents can be used for the
 CC treatment and prevention of diseases such as chronic and acute
 CC inflammation, arthritis, septicaemia, autoimmune diseases, transplant
 CC rejection, stroke, cancer, Alzheimer's disease.
 SQ Sequence 300 AA;

Query Match 100.0%; Score 841; DB 1; Length 300;
 Best Local Similarity 100.0%; Pred. No. 3.4e-66;
 Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LERCRCYCNVLCGEREEARACHATHNRACRTGFFAHAGFCLEHASCPGAGVIAPGTP 60
 |||||
 DB 85 LERCRCYCNVLCGEREEARACHATHNRACRTGFFAHAGFCLEHASCPGAGVIAPGTP 144
 |||||

QY 61 SNTQOCPCPPGTFSSSSSQCOQHNRCTALGLNVPSSSHDTLCSTCGFFPLSTR 120
 |||||
 DB 145 SNTQOCPCPPGTFSSSSSQCOQHNRCTALGLNVPSSSHDTLCSTCGFFPLSTR 204
 |||||

QY 121 VPGAECERAVIDFVAFQDISIKRLQRLQALE 153
 |||||
 DB 205 VPGAECERAVIDFVAFQDISIKRLQRLQALE 237
 |||||

RESULT 3
 W63623

ID W63623 standard; Protein; 170 AA.
 AC W63623;
 DT 26-OCT-1998 (first entry)
 DE Human tumour necrosis factor receptor-6 beta protein.
 KW Human tumour necrosis factor receptor-6 alpha; TNFR-6 alpha; TNFR-6 beta;
 KW endothelial cells; keratinocytes; normal prostate; apoptosis;
 KW prostate tumour tissue.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Peptide 1..30
 FT Protein 31..170
 FT note="TNFR-6 beta"
 FT Region 31..166
 FT /note="Soluble extracellular domain"

PN W09830694-A2.
 PD 16-JUL-1998.
 PE 13-JAN-1998; U00153.
 PR 14-JAN-1997; US-035496.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ebner R, Feng P, Gentz RL, Ni J, Ruben SM, Yu G;
 DR WPI: 98-399142/34.
 DR N-PSDB: V39086.
 PT Human tumour necrosis factor receptors 6-alpha and 6-beta - used in
 PT the diagnosis of immune system-related disorder(s)

PS Claim 20; Fig 2A-2B; 91pp; English.
 CC The present sequence represents the human tumour necrosis factor
 CC receptor-6 beta (TNFR-6 beta) protein. The invention also provides
 CC for the TNFR-6 alpha protein (W63623). TNFR-6 alpha and TNFR-6 beta are
 CC members of the tumour necrosis factor receptor (TNFR) family. TNFRs
 CC are expressed in endothelial cells, keratinocytes, normal prostate and
 CC prostate tumour tissue. For a number of disorders of these cells,
 CC particularly of the immune system, substantially altered (whether
 CC increased or decreased) levels of TNFR-6 alpha and/or TNFR-6 beta gene
 CC expression can be detected, therefore the TNFR-6 alpha and TNFR-6 beta
 CC polypeptides, nucleic acids and antibodies are claimed to be useful in
 CC the diagnosis of such disorders. Mutations of the TNFR-6 alpha and
 CC TNFR-6 beta genes can also be detected. The TNFR polypeptides are
 CC also claimed to be useful for identifying ligands which may be useful
 CC in the treatment of apoptosis related disorders.
 SQ Sequence 170 AA;

Query Match 40.1%; Score 337; DB 1; Length 170;
 Best Local Similarity 100.0%; Pred. No. 1.4e-22;
 Matches 58; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 LERCRCYCNVLCGEREEARACHATHNRACRTGFFAHAGFCLEHASCPGAGVIAPG 58
 |||||
 DB 85 LERCRCYCNVLCGEREEARACHATHNRACRTGFFAHAGFCLEHASCPGAGVIAPG 142
 |||||

RESULT 4
 R99940

ID R99940 standard; Protein; 326 AA.
 AC R99940;
 DT 23-APR-1997 (first entry)
 DE Mutated OCIF, OCIF-DDD1.
 KW Osteoclastogenesis inhibitory factor; OCIF; heparin; bone resorption;
 KW osteoporosis.
 OS Synthetic.
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT protein 22..326
 FT /note="Signal peptide"
 FT /note="Mature OCIF-DDD1"
 FT misc_difference 198..199
 FT /note="Position of deletion, delta 178-252"

PN W09626217-A1.
 PD 29-AUG-1996.
 PF 20-FEB-1996; J00374.
 PR 20-FEB-1995; JP-054977.
 PR 21-JUL-1995; JP-207508.
 PA (SNOW) SNOW BRAND MILK PROD CO LTD.
 PI Goto M, Higashio K, Kobayashi F, Mochizuki S, Morinaga T;
 PI Nakagawa N, Shima N, Tsuda E, Ueda M, Yano K, Yasuda H;
 DR WPI: 96-402320/40.
 DR N-PSDB: T33170.
 PT DNA encoding osteoclastogenesis inhibitory factor protein - useful
 PT for bone resorption control, esp. treatment of osteoporosis
 PS Claim 56; Page 113-114; 183pp; Japanese.
 CC This sequence represents a mutated version of the full length
 CC osteoclastogenesis inhibitory factor (OCIF) of the invention. This
 CC sequence represents OCIF-DDD1 in which amino acids 178-252 of the
 CC mature OCIF protein are deleted. The OCIF of the invention
 CC has a molecular weight by SDS-PAGE of 60 kD under reducing conditions
 CC and 120 kD under non-reducing conditions. The protein is adsorbed onto
 CC cation-exchangers or heparin and its activity is lowered after 10 mins
 CC at 70 deg.C or 30 mins at 56 deg.C, and is lost after 10 mins at 90
 CC deg.C. OCIF is useful in the control of bone resorption and therefore
 CC in the treatment and prevention of disorders of bone resorption, e.g.
 CC osteoporosis.
 SQ Sequence 326 AA;

Query Match 38.6%; Score 324.5; DB 1; Length 326;
 Best Local Similarity 37.9%; Pred. No. 3.5e-21;
 Matches 58; Conservative 33; Mismatches 57; Indels 5; Gaps 2;

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 20:59:03 ; Search time 33.52 seconds
(without alignments)
170.985 Million cell updates/sec

Title: US-09-212-270-2

Perfect score: 841

Sequence: 1 LERCRCYNLCGERBEARA.....FVAFQDISIKRLQLLALE 153

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database : PIR_58:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|--------|--------------------|
| 1 | 241 | 28.7 | 461 | 2 | A35356 | tumor necrosis fac |
| 2 | 221.5 | 26.3 | 459 | 2 | I48854 | gene murine tumour |
| 3 | 218.5 | 26.0 | 474 | 2 | B38634 | tumor necrosis fac |
| 4 | 176.5 | 21.0 | 349 | 2 | D36858 | gene G4R protein - |
| 5 | 176 | 20.9 | 435 | 2 | I54182 | tumor necrosis fac |
| 6 | 157.5 | 18.7 | 325 | 2 | B43692 | T2 protein - rabbl |
| 7 | 154.5 | 18.4 | 277 | 2 | A60771 | B-cell activation |
| 8 | 149.5 | 17.8 | 272 | 2 | I37552 | Ox40 homologue - h |
| 9 | 148 | 17.6 | 289 | 2 | A46515 | B cell-associated |
| 10 | 148 | 17.6 | 305 | 2 | A46476 | CD40 - mouse |
| 11 | 145.5 | 17.3 | 271 | 2 | S12783 | OX40 antigen precu |
| 12 | 140.5 | 16.7 | 416 | 2 | JN0006 | nerve growth facto |
| 13 | 139.5 | 16.6 | 595 | 2 | A42086 | CD30 antigen precu |
| 14 | 137.5 | 16.3 | 272 | 2 | I48700 | gene ox40 protein |
| 15 | 131 | 15.6 | 425 | 2 | A26431 | nerve growth facto |
| 16 | 123.5 | 14.7 | 326 | 1 | GOV2ML | T2 protein - myxom |
| 17 | 123.5 | 14.7 | 427 | 1 | GOHJUN | nerve growth facto |
| 18 | 123 | 14.6 | 256 | 2 | B32393 | T-cell antigen 4-1 |
| 19 | 117 | 13.9 | 255 | 2 | I38426 | 4-1BB - human |
| 20 | 116 | 13.8 | 255 | 2 | JT0752 | lymphocyte activat |
| 21 | 106.5 | 12.7 | 461 | 1 | GQRTT1 | tumor necrosis fac |
| 22 | 105.5 | 12.5 | 327 | 2 | A46484 | apoptosis-mediati |
| 23 | 102.5 | 12.2 | 454 | 1 | GQNST1 | tumor necrosis fac |
| 24 | 102.5 | 12.2 | 454 | 2 | I57826 | tumor necrosis fac |
| 25 | 102 | 12.1 | 780 | 2 | A34102 | von Willebrand fac |
| 26 | 102 | 12.1 | 335 | 2 | A40036 | apoptosis-mediati |
| 27 | 101.5 | 12.1 | 899 | 2 | G02428 | prohormone convert |
| 28 | 101.5 | 12.1 | 915 | 2 | JC6148 | subtilisin-like pr |
| 29 | 101.5 | 12.1 | 2321 | 2 | S78549 | Notch3 protein - h |
| 30 | 101 | 12.0 | 969 | 2 | A39490 | serine proteinase |
| 31 | 101 | 12.0 | 975 | 2 | JC5570 | subtilisin-like pr |
| 32 | 101 | 12.0 | 962 | 2 | JC5571 | subtilisin-like pr |
| 33 | 100 | 11.9 | 642 | 2 | S52111 | uromodulin precurs |
| 34 | 99.5 | 11.8 | 1607 | 1 | MMWSB2 | laminin gamma-1 ch |
| 35 | 98.5 | 11.7 | 328 | 2 | S00530 | surface antigen CR |
| 36 | 98 | 11.7 | 1680 | 2 | A43434 | furin (pC 3.4.21.7 |
| 37 | 97.5 | 11.6 | 1609 | 1 | MMHUB2 | laminin gamma-1 ch |
| 38 | 97 | 11.5 | 314 | 2 | I37383 | FAS soluble protei |
| 39 | 96.5 | 11.5 | 915 | 2 | B48225 | probable proteol |

| | | | | | | |
|----|------|------|------|---|--------|--------------------|
| 40 | 96 | 11.4 | 2482 | 2 | A47648 | perlecan homolog u |
| 41 | 96 | 11.4 | 2277 | 2 | B47648 | perlecan homolog u |
| 42 | 96 | 11.4 | 1160 | 2 | C47648 | perlecan homolog u |
| 43 | 96 | 11.4 | 1713 | 2 | A53347 | adhesive ligand ep |
| 44 | 95.5 | 11.4 | 915 | 2 | A48225 | subtilisin-like pr |
| 45 | 95.5 | 11.4 | 1548 | 2 | S34583 | serine proteinase |

RESULT 1

A35356
tumor necrosis factor receptor type 2 precursor - human
N:Alternate names: 75K tumor necrosis factor receptor
C:Species: Homo sapiens (man)
C>Date: 14-Sep-1990 #sequence_revision 14-Sep-1990 #text_change 20-Mar-1998
C:Accession: A35356; A36475; A48416; A36007; A23666; B35010; I38094
R:Smith, C.A.; Davis, T.; Anderson, D.; Solam, L.; Beckmann, M.P.; Jerzy, R.; Dower, Science 248, 1019-1023, 1990
A:Title: A receptor for tumor necrosis factor defines an unusual family of cellular a
A:Reference number: A35356; MUID:90260639
A:Accession: A35356
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-461 <SMI>
A:Cross-references: GB:M32315; NID:gl89185; PID:gl89186
R:Kohn, T.; Brewer, M.T.; Baker, S.L.; Schwartz, P.E.; King, M.W.; Hale, K.K.; Squir Proc. Natl. Acad. Sci. U.S.A. 87, 8331-8335, 1990
A:Title: A second tumor necrosis factor receptor gene product can shed a naturally oc
A:Reference number: A36475; MUID:91045991
A:Accession: A36475
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-195, 'R', 197-461 <KOH>
A:Cross-references: GB:M55994; GB:M38549; NID:g339757; PID:g339758
R:Dembic, Z.; Loetscher, H.; Gubler, U.; Pan, Y.C.; Lahm, H.W.; Gentz, R.; Brockhaus, Cytokine 2, 231-237, 1990
A:Title: Two human TNF receptors have similar extracellular, but distinct intracellular
A:Reference number: A48416; MUID:91370690
A:Accession: A48416
A:Status: preliminary
A:Molecule type: protein
A:Residues: 23-461 <DEM>
A:Cross-references: GB:S63368; NID:g235648; PID:g235649
A:Note: sequence extracted from NCBI backbone (NCBIN:63368, NCBIP:63371)
R:Heller, R.A.; Song, K.; Onasch, M.A.; Fischer, W.H.; Chang, D.; Ringold, G.M. Proc. Natl. Acad. Sci. U.S.A. 87, 6151-6155, 1990
A:Title: Complementary DNA cloning of a receptor for tumor necrosis factor and demons
A:Reference number: A36007; MUID:90349572
A:Accession: A36007
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 116-140, 'P', 142-195, 'R', 197-362, 'T', 364-461 <HEL>
A:Cross-references: GB:M35857; NID:g339751; PID:g339752
R:Loetscher, H.; Schlaeger, E.J.; Lahm, H.W.; Pan, Y.C.E.; Lesslauer, W.; Brockhaus, J. Biol. Chem. 265, 20131-20138, 1990
A:Title: Purification and partial amino acid sequence analysis of two distinct tumor
A:Reference number: A23666; MUID:91056048
A:Accession: A23666
A:Status: preliminary
A:Molecule type: protein
A:Residues: 23-40; 65-69; 136-141; 300-306 <LOE>
R:Engelmann, H.; Novick, D.; Wallach, D. J. Biol. Chem. 265, 1531-1536, 1990
A:Title: Two tumor necrosis factor-binding proteins purified from human urine. Eviden
A:Reference number: A35010; MUID:90110215
A:Accession: B35010
A:Status: preliminary
A:Molecule type: protein
A:Residues: 27-31 <ENG>
R:Kuhnert, P.; Kemper, O.; Wallach, D. Gene 150, 381-386, 1994

A:Title: Cloning, sequencing and partial functional characterization of the 5' region of
 A:Reference number: I38094; MUID:95121934

A:Accession: I38094

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-37 <RES>

A:Cross-references: EMBL:X80021; NID:g666044; PID:g825701

C:Genetics:

A:Gene: GDB:TNFR2

A:Cross-references: GDB:I25914; OMIM:191191

A:Map position: lp36.2-1p36.2

A:Introns: 26/3

A:Note: the list of introns is incomplete

C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology

C:Keywords: duplication; glycoprotein; receptor; transmembrane protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-416/Product: tumor necrosis factor receptor type 2 #status experimental <MAT>

F:40-76/Domain: NGF receptor repeat homology <NG1>

F:78-119/Domain: NGF receptor repeat homology <NG2>

F:120-162/Domain: NGF receptor repeat homology <NG3>

F:164-201/Domain: NGF receptor repeat homology <NG4>

F:262-279/Domain: transmembrane #status predicted <TMN>

F:280-461/Domain: intracellular #status predicted <INT>

F:171,193/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 28.78; Score 241; DB 2; Length 461;
 Best Local Similarity 36.28; Pred. No. 1.9e-13;
 Matches 46; Conservative 20; Mismatches 49; Indels 12; Gaps 3;

Qy 4 CRYCNVLCGEREEARACHATHNACRRTGFFAHG-----FCLHASCPPGAGVIAP 57

Db 93 CLSCGSRSSDQVETQACTRQNRICTRPGWYCALSKQECRLCAPLKRPGFGVAP 152

Qy 58 GTPSQNTQCPCPPGTFSSASSSEQCQPHRNCATGALGNVPGSSSHDTLCTCTGFP 117

Db 153 GTERSDVVKCPACPGTFTSNTTSDICRPHQICNVVA-----IPGNASMDAVCTTST-PT 206

Qy 118 STRVPGA 124

Db 207 RSMAPGA 213

RESULT 2

I48854

Gene murine tumour necrosis factor receptor 2 protein - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 23-Feb-1997

C:Accession: I48854

R:Powell, E.E.; Wicker, L.S.; Peterson, L.B.; Todd, J.A.

A:Title: Allelic variation of the type 2 tumor necrosis factor receptor gene.

A:Reference number: I48854; MUID:95178848

A:Accession: I48854

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-459 <RES>

A:Cross-references: EMBL:X76401; NID:g433830; PID:g433831

C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology

Query Match 26.38; Score 221.5; DB 2; Length 459;
 Best Local Similarity 37.28; Pred. No. 6.9e-12;
 Matches 42; Conservative 19; Mismatches 41; Indels 11; Gaps 3;

Qy 4 CRYCNVLCGEREEARACHATHNACRRTGFF----AHAGF---CLEHASCPPGAGVIA 56

Db 79 CLSCSSSCTDQVETRACTKQNRVCAACEAGRYCALKTHSGCRQCMRLSKCGFGFVAS 138

Qy 57 PGTPSQNTQCPCPPGTFSSASSSEQCQPHRNCATGALGNVPGSSSHDTLC 109

Db 139 SRAPNGNVLCACAPGTFSDTTSSTDCRPHRICSILA-----IPGNASTDAVC 187

RESULT 3

B38634

tumor necrosis factor receptor type 2 precursor - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 08-Sep-1997

C:Accession: B38634; A40254; S54816

R:Lewis, M.; Tartaglia, L.A.; Lee, A.; Bennett, G.L.; Rice, G.C.; Wong, G.H.W.; Chen, Proc. Natl. Acad. Sci. U.S.A. 88, 2830-2834, 1991

A:Title: Cloning and expression of cDNAs for two distinct murine tumor necrosis facto

A:Reference number: A38634; MUID:91187885

A:Accession: B38634

A:Molecule type: mRNA

A:Residues: 1-474 <LEW>

A:Cross-references: GB:M60469; NID:g199827; PID:g199828

R:Goodwin, R.G.; Anderson, D.; Jerzy, R.; Davis, T.; Brannan, C.I.; Copeland, N.G.; J Mol. Cell. Biol. 11, 3020-3026, 1991

A:Title: Molecular cloning and expression of the type 1 and type 2 murine receptors f

A:Reference number: A40254; MUID:91246188

A:Accession: A40254

A:Molecule type: mRNA

A:Residues: 1-474 <GOO>

A:Cross-references: GB:M60469; NID:g199827; PID:g199828

R:Kisssoneghis, M.; Fellowes, R.; Feldmann, M.; Chernajovsky, Y.

submitted to the EMBL Data Library, May 1995

A:Description: Characterization of the promoter region of the murine p75-TNF receptor

A:Reference number: S54816

A:Accession: S54816

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-22 <KIS>

A:Cross-references: EMBL:X87128; NID:g809043; PID:g809044

C:Superfamily: tumor necrosis factor receptor type 2; NGF receptor repeat homology

C:Keywords: cytokine receptor; transmembrane protein

F:1-22/Domain: signal sequence #status predicted <SIG>

F:23-474/Product: tumor necrosis factor receptor type 2 #status predicted <MAT>

F:40-77/Domain: NGF receptor repeat homology <NG1>

F:79-120/Domain: NGF receptor repeat homology <NG2>

F:166-203/Domain: NGF receptor repeat homology <NG4>

Query Match 26.0%; Score 218.5; DB 2; Length 474;
 Best Local Similarity 37.2%; Pred. No. 1.3e-11;
 Matches 42; Conservative 19; Mismatches 41; Indels 11; Gaps 3;

Qy 4 CRYCNVLCGEREEARACHATHNACRRTGFF----AHAGF---CLEHASCPPGAGVIA 56

Db 94 CLSCSSSCTDQVETRACTKQNRVCAACEAGRYCALKTHSGCRQCMRLSKCGFGFVAS 153

Qy 57 PGTPSQNTQCPCPPGTFSSASSSEQCQPHRNCATGALGNVPGSSSHDTLC 109

Db 154 SRAPNGNVLCACAPGTFSDTTSSTDCRPHRICSILA-----IPGNASTDAVC 202

RESULT 4

D36858

gene G4R protein - variola virus

N:Alternate names: B28R protein (COP)

C:Species: variola virus

C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 10-Sep-1997

C:Accession: D36858; S46888; S32385; S35987

R:Blinov, V.M.

submitted to GenBank, November 1992

A:Description: not shown.

A:Reference number: A36859

A:Accession: D36858

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-349 <BLI>

A:Cross-references: GB:X69198; NID:g456758; PID:g457087

A:Experimental source: strain India-1967, ssp. major, isolate Ind3

R:Kolykhalov, A.A.; Blinov, V.M.; Gytarov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; F

submitted to the EMBL Data Library, April 1992

[illegible]

```
Query Match
Best Local Similarity 18.4%; Score 154.5; DB 2; Length 277;
Matches 34; Conservative 17; Mismatches 52; Indels 5; Gaps 2;

QY 5 KYCNDLVGGEREEARACHATHNRCRCRTGFFAFA---GFCLEHASCPPGAGVIAPGTPS 61
DB 81 KYCNDLVGGEREEARACHATHNRCRCRTGFFAFA---GFCLEHASCPPGAGVIAPGTPS 138

QY 62 QNTQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
DB 139 SDTICEPCPVGFPSNVSAFCKHPWTSCETKDLVVOAGTKNTDVC 186

RESULT 8
OX40 homologue - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: I37552
R:Latza, U.; Durkop, H.; Schnittger, S.; Ringeling, J.; Eitelbach, F.; Hummel, M.; Fonat
Eur. J. Immunol. 24, 677-683, 1994
A:Title: The human OX40 homologue: cDNA structure, expression and chromosomal assignment
A:Reference number: I37552; MUID:94170844
A:Accession: I37552
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <RES>
A:Cross-references: EMBL:X75962; NID:g472957; PID:g472958

Query Match
Best Local Similarity 17.8%; Score 149.5; DB 2; Length 277;
Matches 35; Conservative 9; Mismatches 39; Indels 23; Gaps 4;

QY 4 CRYCNVLGGEREEARACHATHNRCRCRTGFFAFAHAGFLEHASCPPGAGVIAPGTPSON 63
DB 84 CWCNLRSG--SERKOLCFATQTVCRKAG-----TQPLDSYKPG-----123

QY 64 TQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
DB 123 VDCAPCPGPHF--SPGDNQACKPWTNCTLAGKHTLQPA NSSDAIC 166

RESULT 9
A46515
B cell-associated surface molecule CD40 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A46515
R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockayne,
J. Immunol. 149, 3921-3926, 1992
A:Title: Genomic structure and chromosomal mapping of the murine CD40 gene.
A:Reference number: A46515; MUID:93094586
A:Accession: A46515
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-289 <GRI>
A:Experimental source: BALB/c, liver
A:Note: sequence extracted from NCBI backbone (NCBIP:120357)

Query Match
Best Local Similarity 17.6%; Score 148; DB 2; Length 289;
Matches 36; Conservative 15; Mismatches 54; Indels 8; Gaps 3;

QY 3 RC---RYCNVLGGEREEARACHATHNRCRCRTGFFA---HAGFLEHASCPPGAGVIA 56
DB 76 RCHQHRHCEPNQGLRVKKEGT--AESDTVCTCKEGHCTSKDCEACAQHTPCIPGFGYME 133

QY 57 PGTPSONTCQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
DB 134 MATETDTVCHPCVPVGFPSNQSSLFCKYPTWTSCEDKNLEVLQKGTSTQTNVIC 186

Query Match
Best Local Similarity 17.6%; Score 148; DB 2; Length 305;
Matches 36; Conservative 15; Mismatches 54; Indels 8; Gaps 3;

QY 3 RC---RYCNVLGGEREEARACHATHNRCRCRTGFFA---HAGFLEHASCPPGAGVIA 56
DB 76 RCHQHRHCEPNQGLRVKKEGT--AESDTVCTCKEGHCTSKDCEACAQHTPCIPGFGYME 133

QY 57 PGTPSONTCQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
DB 134 MATETDTVCHPCVPVGFPSNQSSLFCKYPTWTSCEDKNLEVLQKGTSTQTNVIC 186

RESULT 11
SI2783
OX40 antigen precursor - rat
N:Alternate names: nerve growth factor receptor homolog
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 10-Sep-1997
C:Accession: SI2783; S08036
R:Mallett, S.; Fossum, S.; Barclay, A.N.
EMBO J. 9, 1063-1068, 1990
A:Title: Characterization of the MRC OX40 antigen of activated CD4 positive T lymphoc
A:Reference number: SI2783; MUID:90214614
A:Accession: SI2783
A:Molecule type: mRNA
A:Residues: 1-271 <MAL>
A:Cross-references: EMBL:X17037; NID:g57830; PID:g57831
C:Keywords: growth factor receptor; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-271/Product: OX40 antigen #status predicted <MAT>
F:211-235/Domain: transmembrane #status predicted <TM>

Query Match
Best Local Similarity 17.3%; Score 145.5; DB 2; Length 271;
Matches 36; Conservative 13; Mismatches 31; Indels 35; Gaps 5;

QY 2 ECRYCNVLGGEREEARACHATHNRCRCRTGFFAFAHAGFLEHASCPPGAGVIAPGT-P 60
DB 77 KOCTOCNHRSG--SELKONCTPTEDTVCCOR-----PGTQP 110

QY 61 SONT-----QCQCPGPGTFSASSSSSEQCPHRNCTALGLALNVPGSSSHDTLC 109
DB 111 RODSSHKLGVDCVPCPGPHF--SPGSNQACKPWTNCTLSGKQIRHPASNSLDTVC_163

RESULT 12
JN0006
nerve growth factor receptor, low affinity - chicken
N:Alternate names: NGF receptor
C:Species: Gallus gallus (chicken)
```

C>Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 01-May-1998
C:Accession: JN0006; A60504
R:Large, T.H.; Weskamp, G.; Helder, J.C.; Radeke, M.J.; Misko, T.P.; Shooter, E.M.; Reid
Neuron 2, 1123-1134, 1989
A:Title: Structure and developmental expression of the nerve growth factor receptor in t
A:Reference number: JN0006; MUID:90166579
A:Accession: JN0006
A:Molecule type: mRNA
A:Residues: 1-416 <LAR>
A:Experimental source: embryonic chick brain
R:Heuer, J.G.; Fatemie-Nainie, S.; Wheeler, E.F.; Bothwell, M.
Dev. Biol. 137, 287-304, 1990
A:Title: Structure and developmental expression of the chicken NGF receptor.
A:Reference number: A60504
A:Accession: A60504
A>Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 21-35, Y, 37-172, 'K', 174-275, 'S', 277-395, 'R', 397-416 <HEU>
C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma
C:Comment: The cysteine-rich region of the extracellular domain may form part or all of
C:Comment: This protein is thought to form a high-affinity receptor when it associates
C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor; t
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-416/Product: nerve growth factor receptor #status predicted <MAT>
F:21-239/Domain: extracellular #status predicted <EXT>
F:24-57/Domain: NGF receptor repeat homology <NG1>
F:59-100/Domain: NGF receptor repeat homology <NG2>
F:101-139/Domain: NGF receptor repeat homology <NG3>
F:141-181/Domain: NGF receptor repeat homology <NG4>
F:189-237/Region: serine/threonine-rich
F:240-261/Domain: transmembrane #status predicted <MEM>
F:262-416/Domain: intracellular #status predicted <INT>
F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.7%; Score 140.5; DB 2; Length 416;
Best Local Similarity 29.8%; Pred. No. 4.6e-05;
Matches 34; Conservative 16; Mismatches 49; Indels 15; Gaps 4;

Qy 2 ECRVCNVLCGEREEARACHATHNRACRRTGTFPAH--AGFCLEHASCPPGAGVIAPGT 59
Db 73 EPCKPCTQCVGLHMSA-PCVESDDAVRCAYGFQDELSGSKSCSICEVGFGLMFPGR 131

Qy 60 PSQNTCCQPCPTFSASSSSSQCPHNRCTALGLALNVPGSSSHDTLCSTCT 113
Db 132 DSQDVTCECPGTSDANFVDCPLP---CFI-----CEENEVNVKECT 173

RESULT 13
A42086
CD30 antigen precursor - human
N:Alternate names: Ki-1 antigen; nerve growth factor receptor family member CD30
C:Species: Homo sapiens (man)
C>Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 24-Sep-1998
C:Accession: A42086
R:Durkop, H.; Latza, U.; Hummel, M.; Eitelbach, F.; Seed, B.; Stein, H.
Cell 68, 421-427, 1992
A:Title: Molecular cloning and expression of a new member of the nerve growth factor rec
A:Reference number: A42086; MUID:92154659
A:Accession: A42086
A:Molecule type: mRNA
A:Residues: 1-595 <DUR>
A:Cross-references: GB:M83554; NID:g180095; PID:g180096
A:Experimental source: HUT-102 cell line
A:Note: sequence extracted from NCBI backbone (NCBI:82088, NCBI:82090)
C:Genetics:
A:Gene: GDB:CD30; D1S166E
A:Cross-references: GDB:131547; OMIM:153243
A:Map position: 1p36-1p36
C:Superfamily: NGF receptor repeat homology
C:Keywords: glycoprotein; growth factor receptor; transmembrane protein
F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-383/Domain: extracellular #status predicted <EXT>
F:384-407/Domain: transmembrane #status predicted <TMM>
F:408-595/Domain: intracellular #status predicted <CYT>
F:101,276/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 16.6%; Score 139.5; DB 2; Length 595;
Best Local Similarity 31.4%; Pred. No. 7.7e-05;
Matches 38; Conservative 12; Mismatches 50; Indels 21; Gaps 6;

Qy 2 ECRVCNVLCGEREEARACHATHNRACRRTGTFPAHGF-----CLEHASCPPGAGV 54
Db 79 DRCTAC-VTCSRDDLVEKTPCAWNSRVCECRPGMFCSTSAVNSCARCFHSHVCPAGMIV 137

Qy 55 IAPGTPSONTOCQPCPTFSASSSSSQCPHNRCTALGLALNVPGSSSHDT 107
Db 138 KFGPTAQKNTVCEPASPGV-SPACASPENCKPESSGTIPQAKPTPVS-----PATSSAST 191

Qy 108 L 108
Db 192 M 192

RESULT 14
I48700
gene ox40 protein - mouse
N:Alternate names: OX40 antigen
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Oct-1996
C:Accession: I48700; I48334; S34377
R:Calderhead, D.M.; Buhlmann, J.E.; van den Eertwegh, A.J.; Claassen, E.; Noelle, R.J
J. Immunol. 151, 5261-5271, 1993
A:Title: Cloning of mouse Ox40: a T cell activation marker that may mediate T-B cell
A:Reference number: I48700; MUID:94044750
A:Accession: I48700
A>Status: translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-272 <RES>
A:Cross-references: EMBL:Z21674; NID:g312827; PID:g312828
R:Birkeland, M.L.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Barclay, A.N.
Eur. J. Immunol. 25, 926-930, 1995
A:Title: Gene structure and chromosomal localization of the mouse homologue of rat OX
A:Reference number: I48334; MUID:95255413
A:Accession: I48334
A>Status: translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-14, 'G', 16-272 <RE2>
A:Cross-references: EMBL:X85214; NID:g732818; PID:g732819
C:Genetics:
A:Gene: ox40
A:Introns: 45/1; 86/1; 122/1; 144/2; 210/1; 250/1

Query Match 16.3%; Score 137.5; DB 2; Length 272;
Best Local Similarity 30.4%; Pred. No. 5.7e-05;
Matches 35; Conservative 11; Mismatches 34; Indels 35; Gaps 5;

Qy 2 ECRVCNVLCGEREEARACHATHNRACRRTGTFPAHAGFCLEHASCPPGAGVIAPGT-P 60
Db 78 KOCTOCNHRSG--SELKQNCPTQDTVCRCR-----PGTOP 111

Qy 61 SQNT-----CQCPCTPTFSASSSSSQCPHNRCTALGLALNVPGSSSHDTLC 109
Db 112 RQDSGYKLGVDVCPGPGHF--SPGNNAQCKPWTNCTLSGKQTRHPASDSLDAVC 164

RESULT 15
A26431
nerve growth factor receptor precursor, low affinity - rat
N:Alternate names: NGF receptor
C:Species: Rattus norvegicus (Norway rat)
C>Date: 19-Nov-1988 #sequence_revision 19-Nov-1988 #text_change 01-May-1998
C:Accession: A26431; PH1229

R:Radeke, M.J.; Misko, T.P.; Hsu, C.; Herzenberg, L.A.; Shooter, E.M.
Nature 325, 593-597, 1987
A:Title: Gene transfer and molecular cloning of the rat nerve growth factor receptor.
A:Reference number: A26431; MUID:87115859
A:Accession: A26431
A:Molecule type: mRNA
A:Residues: 1-425 <RAD>
A:Cross-references: GB:X05137; NID:g56755; PID:g56756
R:Metsis, M.; Timmusk, T.; Allikmets, R.; Saarma, M.; Persson, H.
Gene 121, 247-254, 1992
A:Title: Regulatory elements and transcriptional regulation by testosterone and retinoic acid.
A:Reference number: PH1229
A:Accession: PH1229
A:Molecule type: DNA
A:Residues: 1-20 <MET>
A:Cross-references: GB:X61269
C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma cells, and on the cysteine-rich region of the extracellular domain may form part or all of a high-affinity receptor when it associates with a specific ligand.
C:Genetics:
A:Introns: 20/3
C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor; tyrosine kinase
F:1-29/Domain: signal sequence #status predicted <SIG>
F:30-425/Product: nerve growth factor receptor #status predicted <MAT>
F:30-251/Domain: extracellular #status predicted <EXT>
F:33-66/Domain: NGF receptor repeat homology <NG1>
F:68-109/Domain: NGF receptor repeat homology <NG2>
F:110-148/Domain: NGF receptor repeat homology <NG3>
F:150-190/Domain: NGF receptor repeat homology <NG4>
F:198-249/Region: serine/threonine-rich
F:252-273/Domain: transmembrane #status predicted <MEM>
F:274-425/Domain: intracellular #status predicted <INT>
F:61/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 15.6%; Score 131; DB 2; Length 425;
Best Local Similarity 27.3%; Pred. No. 0.0003;
Matches 36; Conservative 15; Mismatches 53; Indels 28; Gaps 5;
QY 2 ECRYCNVLCGEREEPARACHATHNRACRTGFF--AHAGFCLEHASCPPGAGVIAPGT 59
Db 82 EPKPCTECLGLQMSA-PCVEADDAVCRCAVGYODEETHGHCACSVCEVSGGLVFSQ 140
QY 60 PSQNTCCQPCPTFFASSSSSFOCPHNCALGLALNVPSSSHDYLCTSC--TGFP 117
Db 141 DRQNTVCECPGTYSDANHVDPCLP-----CTVCEPTEROL 178
QY 118 STRVGAEE-ECE 128
Db 179 RECTPWADAECE 190

Search completed: May 13, 1999, 20:59:04
Job time: 2318 sec

RA ZHENG Z., NAISMITH J.H., THOMAS D.;
RL PROTEIN SCI. 7:1124-1135(1998).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: X60592; E1175755; -.
DR PIR: S04460; S04460.
DR PDB: 1CDF; 01-APR-97.
DR MM: 109535; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 1.
DR PROSITE: PS50050; TNFR_NGFR_2; 4.
KM RECEPTOR: B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL;
KW 3D-STRUCTURE.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 277 CD40L RECEPTOR.
FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 POTENTIAL.
FT DOMAIN 216 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT DISULFID 26 37
FT DISULFID 38 51
FT DISULFID 41 59
FT DISULFID 62 77
FT DISULFID 83 103
FT DISULFID 105 119
FT DISULFID 111 116
FT DISULFID 125 143
FT CARBOHYD 153 153 POTENTIAL.
FT CARBOHYD 180 180 POTENTIAL.
SQ SEQUENCE 277 AA; 30619 MM; 38284411 CRC32;

Query Match 18.48; Score 154.5; DB 1; Length 277;
Best Local Similarity 31.5%; Pred. No. 1.7e-06;
Matches 34; Conservative 17; Mismatches 52; Indels 5; Gaps 2;
OY 5 RYCNVLCGEREERERACHATHNACRCRTGFFAHA---GFCLEHASCPPGAGVIAPGTPS 61
DB 81 KYCDNGLRLRVOQKTSST--DTICTCEGWHCTSEACECVLHRSCTSPGFGVKQIANGV 138
OY 62 ONTCOCPPPGPTFSASSSSSEQCOPHRNCTALGLALNVPSSSHDTLC 109
DB 139 SDTICEPCPGVGFSSNVSSAFECHEPWTSCETKDLVVOAGTKTKTVVC 186

RESULT 8
OX40_HUMAN STANDARD; PRT: 277 AA.
ID OX40_HUMAN
AC P43489; 013663;
DT 01-NOV-1995 (REL. 32, CREATED)
DT 01-NOV-1995 (REL. 32, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE OX40L RECEPTOR PRECURSOR (ACT35 ANTIGEN) (TAX-TRANSCRIPTIONALLY
DE ACTIVATED GLYCOPROTEIN 1 RECEPTOR) (CD134 ANTIGEN).
GN TXGPII.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 94170844.
RA LAZDA U., DUTERKOP H., SCHNITTEGER S., RINGELING J., EITELBACH F.,
RA HUMMEL M., FONATSCHE C., STEIN H.;
RL EUR. J. IMMUNOL. 24:677-683(1994).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95219871.
RA BAUM P.R., GAILE R.B. III, RAMSDALL F., SRINIVASAN S., SORESENSEN R.A.,

RA WATSON M.L., SELDIN M.F., CLIFFORD K.N., GRABSTEIN K., ALDERSON M.R.;
RL CIRC. SHOCK 44:30-34(1994).
CC -1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: X75962; G472958; -.
DR EMBL: S76792; G913406; ALT_INIT.
DR MM: 600315; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 3.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KM RECEPTOR: T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
KW SIGNAL.
FT SIGNAL 1 28 POTENTIAL.
FT CHAIN 29 277 OX40L RECEPTOR.
FT DOMAIN 29 214 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 215 235 POTENTIAL.
FT DOMAIN 236 277 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 30 169 4 X TNFR-CYS.
FT REPEAT 30 65 TNFR-CYS 1.
FT REPEAT 66 107 TNFR-CYS 2.
FT REPEAT 108 126 TNFR-CYS 3 (INCOMPLETE).
FT REPEAT 127 167 TNFR-CYS 4.
FT CARBOHYD 146 146 POTENTIAL.
FT CARBOHYD 160 160 POTENTIAL.
SQ SEQUENCE 277 AA; 29340 MM; E0BB6765 CRC32;

Query Match 17.88; Score 149.5; DB 1; Length 277;
Best Local Similarity 33.0%; Pred. No. 4.6e-06;
Matches 35; Conservative 9; Mismatches 39; Indels 23; Gaps 4;
OY 4 CRYCNVLCGEREERERACHATHNACRCRTGFFAHAAGFCLEHASCPPGAGVIAPGTPSON 63
DB 84 CTWCMILRSG--SERKQLCTATQDTVCRCRAG-----TOPLDYSKPG----- 123
OY 64 TOCOCPPPGPTFSASSSSSEQCOPHRNCTALGLALNVPSSSHDTLC 109
DB 123 VDCACCPPGHF--SPGDNQACKPWTNCTLAGKHTLQPPASSNSDAIC 166

RESULT 9
CD40_MOUSE STANDARD; PRT: 289 AA.
ID CD40_MOUSE
AC P27512;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDM40).
GN CD40.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92105763.
RA TORRES R.M., CLARK E.A.;
RL J. IMMUNOL. 148:620-626(1992).
RN [2]
RP REVISIONS.
RC STRAIN-BALB/C;
RA TORRES R.M.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE-LIVER;
RX MEDLINE: 93094586.
RA GRIMALDI J.C., TORRES R., KOZAK C.A., CHANG R., CLARK E.A.,
RA HOWARD M., COCKAYNE D.A.;
RL J. IMMUNOL. 149:3921-3926(1992).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: M83312; G1553059; -.

DR EMBL: M94126; G192526; .
 DR EMBL: M94129; G192526; JOINED.
 DR EMBL: M94128; G192526; JOINED.
 DR EMBL: M94127; G192526; JOINED.
 DR PIR: A46476; A46476.
 DR HSP: P19438; 1TNR.
 DR MGD: MGI:88336; CD40.
 DR PROSITE: PS00652; TNR_NGFR_1; 1.
 DR PROSITE: PS0050; TNR_NGFR_2; 4.
 DR RECEPTOR: B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL.
 FT SIGNAL 1 19
 FT CHAIN 20 289
 FT DOMAIN 20 193
 FT TRANSMEM 194 215
 FT DOMAIN 216 289
 FT DOMAIN 25 187
 FT REPEAT 25 60
 FT REPEAT 61 103
 FT REPEAT 104 144
 FT REPEAT 145 187
 FT CARBOHYD 153 153
 FT SEQUENCE 289 AA; 32111 MW; D8D70A2C CRC32;

Query Match 17.6%; Score 148; DB 1; Length 289;
 Best Local Similarity 31.9%; Pred. No. 6.4e-06;
 Matches 36; Conservative 15; Mismatches 54; Indels 8; Gaps 3;

OY 3 RC---HYCNVLGCEEREDEARACHATNHRACRCRTGFFA---HAGFCLHASCPPGAGVIA 56
 DB 76 RCQHSHCEPNOGLRVKKEGT--AESDVTCTCKEGHCHTSKDEACAOHPCCIPGFGVME 133
 OY 57 PGTSPONTCCPCPGTPTFSASSSSSECCOPHRNCTALGLALNVPGSSSHDTLC 109
 DB 134 MATETDTYCHPCVPVGFSSFLFKCYPTWCSEKKNLEVLKGTSGTINVIC 186

RESULT 10
 OX40_RAT STANDARD; PRI: 271 AA.
 ID OX40_RAT
 AC P15725;
 DT 01-APR-1990 (REL. 14, CREATED)
 DT 01-APR-1990 (REL. 14, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN) (MRC OX40).
 GN TXGP1L OR OX40.
 OS RATUUS NORVEGICUS (RAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-T-CELL.
 RX MEDLINE: 90214614.
 RA MALLETT S., FOSSUM S., BARCLAY A.N.;
 RL EMO J. 9:1063-1068(1990).
 CC -1- FUNCTION: RECEPTOR FOR THE OX40L/SP34 CYTOKINE.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: ACTIVATED T-CELLS.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: X17037; G57831; .
 DR PIR: S08036; S08036.
 DR PROSITE: PS00652; TNR_NGFR_1; 3.
 DR PROSITE: PS00652; TNR_NGFR_2; 2.
 DR RECEPTOR: T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
 KW SIGNAL.
 FT SIGNAL 1 19
 FT CHAIN 20 271
 FT DOMAIN 20 210
 FT TRANSMEM 216 235
 FT DOMAIN 236 271
 FT DOMAIN 25 164
 FT REPEAT 60 4 X TNR-CYS.
 FT TNR-CYS 1. 60

FT REPEAT 61 102 TNR-CYS 2.
 FT REPEAT 103 123 TNR-CYS 3 (INCOMPLETE).
 FT REPEAT 124 164 TNR-CYS 4.
 FT CARBOHYD 143 143 POTENTIAL.
 FT SEQUENCE 271 AA; 29895 MW; 65460ED2 CRC32;

Query Match 17.3%; Score 145.5; DB 1; Length 271;
 Best Local Similarity 31.3%; Pred. No. 9.7e-06;
 Matches 36; Conservative 13; Mismatches 31; Indels 35; Gaps 5;

OY 2 ERCRYCNVLGCEEREDEARACHATNHRACRCRTGFFAHAGFCLHASCPPGAGVIA 60
 DB 77 KCTQCNNHSG--SELKONTPTEDTVCOR-----PGTOP 110
 OY 61 SONT-----QCPCPPGTPTFSASSSSSECCOPHRNCTALGLALNVPGSSSHDTLC 109
 DB 111 RDSHRLAGVDCVPCPGHF--SPGSNQACKPWTNCTLSGKQIRHPSNSLDTVC 163

RESULT 11
 NGFR_CHICK STANDARD; PRI: 416 AA.
 ID NGFR_CHICK
 AC P18519;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)
 DE (GP80-LINGFR) (P75 ICD).
 GN NGFR.
 OS GALLUS GALLUS (CHICKEN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
 OC GALLIFORMES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE: 90166579.
 RA LARGE T.H., WESKAMP G., HEIDER J.C., RADEKE M.J., MISKO T.P.,
 RA SHOOTER E.M., REICHARDT L.F.;
 RL NEURON 2:1123-1134(1989).
 RN [2]
 RP SEQUENCE OF 21-416 FROM N.A.
 RX MEDLINE: 90152140.
 RA HEUER J.G., FATEME-NAINE S., WHEELER E.F., BOTHWELL M.;
 RL DEV. BIOL. 137:287-304(1990).
 CC -1- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
 CC NT-3, AND NT-4.
 CC -1- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
 CC BOND FORMATION.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNR-TYPE CYSTEINE-RICH REGION.
 DR PIR: J00006; J00006.
 DR PIR: A60504; A60504.
 DR PROSITE: PS00652; TNR_NGFR_1; 3.
 DR PROSITE: PS00652; TNR_NGFR_2; 3.
 DR PROSITE: PS00017; DEATH_DOMAIN; 1.
 DR RECEPTOR: NEUROGENESIS; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT;
 KW PHOSPHORYLATION; SIGNAL.
 FT SIGNAL 1 19
 FT CHAIN 20 416
 FT DOMAIN 29 239
 FT TRANSMEM 240 261
 FT DOMAIN 262 416
 FT DOMAIN 23 181
 FT REPEAT 23 57
 FT REPEAT 58 100
 FT REPEAT 101 140
 FT REPEAT 141 181
 FT REPEAT 188 236
 FT DOMAIN 333 410
 FT DISULFID 24 35
 FT DISULFID 36 49
 FT BY SIMILARITY.

```
FT DISULFID 39 56 BY SIMILARITY.
FT DISULFID 59 75 BY SIMILARITY.
FT DISULFID 78 91 BY SIMILARITY.
FT DISULFID 81 99 BY SIMILARITY.
FT DISULFID 101 114 BY SIMILARITY.
FT DISULFID 117 130 BY SIMILARITY.
FT DISULFID 120 138 BY SIMILARITY.
FT DISULFID 141 156 BY SIMILARITY.
FT DISULFID 159 172 BY SIMILARITY.
FT DISULFID 162 180 BY SIMILARITY.
FT CARBOHYD 52 52 POTENTIAL.
FT CONFLICT 36 36 C -> Y (IN REF. 2).
FT CONFLICT 173 173 T -> K (IN REF. 2).
FT CONFLICT 276 276 N -> S (IN REF. 2).
FT CONFLICT 396 396 K -> R (IN REF. 2).
SQ SEQUENCE 416 AA: 44654 MW: 4D3F086A CRC32;

Query Match 16.7%; Score 140.5; DB 1; Length 416;
Best Local Similarity 29.8%; Pred. No. 3.8e-05;
Matches 34; Conservative 16; Mismatches 49; Indels 15; Gaps 4;

QY 2 ECRYCNVLGGEREERARACHATNRACRCRTGFFAH--AGFCLERHASCPPGAGVIAPGT 59
Db 73 EPCPKPTQCVGLHSMRA-PCVESDVAVCRAVCYFODELSGCKESICGVGFIAPCR 131
QY 60 PSONTCQPCPPGPTFSASSSSSEOCOPHRNCTALGLALNVPGSSSHDTICTSCT 113
Db 132 DSQDTVCECEPCEGTFSEAFNFDPLP---CTI-----CEENEVMVKECT 173

RESULT 12
CD30_HUMAN
ID CD30_HUMAN STANDARD; PRT; 595 AA.
AC P28908;
DT 01-DEC-1992 (REL. 24, CREATED)
DT 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE CD30L RECEPTOR PRECURSOR (LYMPHOCYTE ACTIVATION ANTIGEN CD30) (KI-1 ANTIGEN).
GN CD30.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LYMPHOID;
RX MEDLINE: 92154659.
RA DUEKOP H., LATZA U., HUMMEL M., EITELBACH F., SEED B., STEIN H.;
RL CELL 68:421-427(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95089787.
RA JUNG W., KUEGER S., RENNER C., GAUSE A., SAHIN U., TRUMPER L.,
RA PERUNDUCHUH M.;
RL MOL. IMMUNOL. 31:1329-1334(1994).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD30L. MAY PLAY
CC A ROLE IN THE REGULATION OF CELLULAR GROWTH AND TRANSFORMATION OF
CC ACTIVATED LYMPHOBLASTS.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- DISASE: MOST SPECIFIC HODGKIN'S DISEASE ASSOCIATED ANTIGEN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD30 entry;
CC WWW="http://www.ncbi.nlm.nih.gov/Prov/cd/cd30.htm".
DR EMBL: M83554; G180096; -.
DR EMBL: S75768; E184151; -.
DR PIR: A42086; A42086.
DR MIM: 153243; -.
DR PROSITE: P500652; TNFR_NGFR_1; 2.
DR PROSITE: P50050; TNFR_NGFR_2; 2.
KW T-CELL; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL; RECEPTOR; REPEAT.
FT SIGNAL 1 18
FT CHAIN 19 595 CD30L RECEPTOR.
```

```
FT DOMAIN 19 379 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 380 407 POTENTIAL.
FT DOMAIN 408 595 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 28 325 6 X TNFR-CYS.
FT REPEAT 28 66 TNFR-CYS 1.
FT REPEAT 68 106 TNFR-CYS 2.
FT REPEAT 107 150 TNFR-CYS 3.
FT REPEAT 205 241 TNFR-CYS 4.
FT REPEAT 243 281 TNFR-CYS 5.
FT REPEAT 282 325 TNFR-CYS 6.
FT DOMAIN 347 377 PRO/SER/THR-RICH.
FT CARBOHYD 101 101 POTENTIAL.
FT CARBOHYD 276 276 POTENTIAL.
SQ SEQUENCE 595 AA: 63747 MW: A46D6C0B CRC32;

Query Match 16.6%; Score 139.5; DB 1; Length 595;
Best Local Similarity 31.4%; Pred. No. 6.4e-05;
Matches 38; Conservative 12; Mismatches 50; Indels 21; Gaps 6;

QY 2 ECRYCNVLGGERE-REARACHATNRACRCRTGFFAHG-----CLEHASCPPGAGV 54
Db 79 DRCTAC-VTCSRDLYEKPICAMNSSRYVCECRGMFCSTSAVNSCARCFHSVCPAGMIV 137
QY 55 IAPGTSQNTQCPGPTFSASSSSSEOC-----PHRNCTALGLALNVPGSSSHDT 107
Db 138 KFGPTAQKNTVCBPASPGV-SPACASPEMCKEPPSGTIIQAKPTPV-----PATSSAST 191
QY 108 L 108
Db 192 M 192

RESULT 13
CD40L_BOVIN
ID CD40L_BOVIN STANDARD; PRT; 269 AA.
AC Q28203;
DT 01-NOV-1997 (REL. 35, CREATED)
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (FRAGMENT).
GN CD40.
OS BOS TAURUS (BOVINE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; ARTIODACTYLA.
RN [1]
RP SEQUENCE FROM N.A.
RA ESTES D.M., HIRANO A.;
RL SUBMITTED (AUG-1996) TO EMBL/GENBANK/DDJ DATA BANKS.
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: U57745; G1480643; -.
DR PROSITE: P500652; TNFR_NGFR_1; 1.
DR PROSITE: P50050; TNFR_NGFR_2; 1.
KW RECEPTOR; B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL.
FT SIGNAL 1 19
FT CHAIN 20 269
FT DOMAIN 20 193 CD40L RECEPTOR.
FT TRANSMEM 194 215 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 216 269 POTENTIAL.
FT DOMAIN 216 269 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT CARBOHYD 153 153 TNFR-CYS 4.
FT CARBOHYD 180 180 POTENTIAL.
FT NON_TER 269
SQ SEQUENCE 269 AA: 29983 MW: 46C3AB4D CRC32;

Query Match 16.3%; Score 137.5; DB 1; Length 269;
```

Best Local Similarity 31.5%; Pred. No. 4.5e-05;
Matches 35; Conservative 11; Mismatches 54; Indels 11; Gaps 2;

OY 5 RYCVVIGEREEREA-----RACHATHNRACRCRTGFAHAGFCLEHASCPRGAGVIAPG 58
DB 81 RYCNPNJGLRIQSGTLTDTICVCEGQHCSTHT-----CESCPHSLCLPGGVQIOA 135
OY 59 TSPONTCCPCPCPTGTFSSSSSECCOPHRNCTALGLALNPVSSSHDTLC 109
DB 136 TGLDTCVCEPPLGFFSNVSSAFKCHWTSCERKGLVEQVGNKTIYVC 166

RESULT 14
OX40_MOUSE STANDARD; PRT; 272 AA.
ID OX40_MOUSE
AC P47741;
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
GN TXGP1 OR OX40.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BALB/C;
RX MEDLINE: 94044750.
RA CALDERHEAD D.M., BUHLMANN J.E., VAN DEN EERTWEGH A.J.,
RA CLAASSEN E., NOELLE R.J., FELL H.;
RL J. IMMUNOL. 151:5261-5271(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 95255413.
RA BIRKELAND M.L., COPELAND N.G., GILBERT D.J., JENKINS N.A.,
RA BARCLAY A.N.;
RL EUR. J. IMMUNOL. 25:926-930(1995).
CC -1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: Z21674; G312828; -;
DR EMBL: X85214; G732819; -;
DR MGD: MGI:104512; TXGP1.
DR PROSITE: PS00652; TNFR_NGFR.1; 3.
DR PROSITE: PS50050; TNFR_NGFR.2; 2.
KW RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
KW SIGNAL.
FT SIGNAL 1 19
FT CHAIN 20 272
FT DOMAIN 20 211
FT TRANSMEM 212 236
FT DOMAIN 237 272
FT DOMAIN 26 165
FT REPEAT 26 61
FT REPEAT 62 103
FT REPEAT 104 124
FT REPEAT 125 165
FT CARBOHYD 144 144
FT CONFLICT 15 15
SQ SEQUENCE 272 AA; 30153 MW; 67D1B978 CRC32;

Query Match 16.3%; Score 137.5; DB 1; Length 272;
Best Local Similarity 30.4%; Pred. No. 4.6e-05;
Matches 35; Conservative 11; Mismatches 34; Indels 35; Gaps 5;

OY 2 ECRATCNVLCGEREEERACHATHNRACRCRTGFAHAGFCLEHASCPRGAGVIAPG-T-P 60
DB 78 KCTQCNRHSG--SELKNCCTPTDTCRCR-----PGTTP 111
OY 61 SONT-----OCQRPPTTFSSSSSECCOPHRNCTALGLALNPVSSSHDTLC 109
DB 112 RODSGYKLGIVDCVPCPGHF--SPGNNOACKPMTNCTLSGKOTRHPASDISDAVC 164

RESULT 15
NGFR_RAT STANDARD; PRT; 425 AA.
ID NGFR_RAT
AC P07174;
DT 01-APR-1988 (REL. 07, CREATED)
DT 01-APR-1988 (REL. 07, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)
DE (GP80-LNGFR) (P75 ICD).
GN NGFR.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87115859.
RA RADEKE M.J., MISKO T.P., HSU C., HERZENBERG L.A., SHOOTER E.M.;
RL NATURE 325:593-597(1987).
RN [2]
RP SEQUENCE OF 1-22 FROM N.A.
RC TISSUE-LIVER.
RX MEDLINE: 93077038.
RA METSIS M., TIMUSK T., ALLIKMETS R., SARMA M., PERSSON H.;
RL GENE 121:247-254(1992).
RN [3]
RP STRUCTURE BY NMR OF 334-418.
RX MEDLINE: 97449145.
RA LIEPINSH E., ILAG L.L., OTTING G., IBANEZ C.F.;
RL EMBO J. 16:4999-5005(1997).
CC -1- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
NT-3, AND NT-4.
CC -1- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
BOND FORMATION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- PFM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: X05137; G65756; -; NOT_ANNOTATED_CDS.
DR PIR: A26431; A26431.
DR PDB: INGR; 29-JUL-97.
DR PROSITE: PS00652; TNFR_NGFR.1; 3.
DR PROSITE: PS50050; TNFR_NGFR.2; 4.
DR PROSITE: PS50017; DEATH_DOMAIN.1.
KW RECEPTOR; NEUROGENESIS; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT;
KW PHOSPHORYLATION; SIGNAL; 3D-STRUCTURE.
FT SIGNAL 1 29
FT CHAIN 30 425
FT DOMAIN 30 251
FT TRANSMEM 252 273
FT DOMAIN 274 425
FT DOMAIN 32 190
FT REPEAT 32 66
FT REPEAT 67 108
FT REPEAT 109 148
FT REPEAT 149 190
FT DOMAIN 198 249
FT DOMAIN 354 419
FT DISULFID 33 44
FT DISULFID 45 58
FT DISULFID 48 65
FT DISULFID 68 84
FT DISULFID 87 100
FT DISULFID 110 123
FT DISULFID 126 139
FT DISULFID 129 147
FT DISULFID 150 165
FT DISULFID 168 181
FT DISULFID 171 189
FT CARBOHYD 71 71
FT POTENTIAL.

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 1999, 21:39:56 ; Search time 48.52 Seconds

(without alignments)
173.967 Million cell updates/sec

Title: US-09-212-270-2

Perfect score: 841

Sequence: 1 LERRRYCNCVLCGEREEREA.....FVAFODISIKRLQRLDALE 153

Scoring table: BLOSUM62

Searched: 180763 segs, 55169189 residues

Database :

SPREMBL_8:*
1: sp_fungi:*
2: sp_human:*
3: sp_invertebrate:*
4: sp_mammal:*
5: sp_mhc:*
6: sp_organelle:*
7: sp_phase:*
8: sp_plant:*
9: sp_bacteria:*
10: sp_rodentia:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|--------|----|--------|--------------------|
| 1 | 309.5 | 36.8 | 401 | 2 | 000300 | 000300 homo sapien |
| 2 | 309.5 | 36.8 | 401 | 2 | 060236 | 060236 homo sapien |
| 3 | 306.5 | 36.4 | 401 | 10 | 008727 | 008727 rattus norv |
| 4 | 303.5 | 36.1 | 401 | 10 | 008712 | 008712 mus musculu |
| 5 | 300.5 | 35.7 | 401 | 10 | 070202 | 070202 mus musculu |
| 6 | 241 | 28.7 | 425 | 2 | Q16042 | Q16042 homo sapien |
| 7 | 221.5 | 26.3 | 459 | 10 | Q62327 | Q62327 mus musculu |
| 8 | 221.5 | 26.3 | 482 | 10 | 088734 | 088734 mus musculu |
| 9 | 192 | 22.4 | 655 | 2 | 075509 | 075509 homo sapien |
| 10 | 188.5 | 22.4 | 348 | 11 | 057277 | 057277 monkeypox v |
| 11 | 188.5 | 22.4 | 348 | 11 | 057103 | 057103 monkeypox v |
| 12 | 188.5 | 22.4 | 348 | 11 | 057108 | 057108 monkeypox v |
| 13 | 184.5 | 21.9 | 360 | 11 | 057118 | 057118 monkeypox v |
| 14 | 184 | 21.9 | 349 | 11 | 057100 | 057100 monkeypox v |
| 15 | 182 | 21.6 | 349 | 11 | 057291 | 057291 monkeypox v |
| 16 | 182 | 21.6 | 349 | 11 | 057101 | 057101 monkeypox v |
| 17 | 182 | 21.6 | 349 | 11 | 057102 | 057102 monkeypox v |
| 18 | 181 | 21.5 | 349 | 11 | 057116 | 057116 monkeypox v |
| 19 | 179.5 | 21.3 | 349 | 11 | 057099 | 057099 monkeypox v |
| 20 | 179.5 | 21.3 | 349 | 11 | 057115 | 057115 monkeypox v |
| 21 | 179.5 | 21.3 | 349 | 11 | 057120 | 057120 monkeypox v |
| 22 | 179.5 | 21.3 | 349 | 11 | 057121 | 057121 monkeypox v |
| 23 | 179.5 | 21.3 | 349 | 11 | 057122 | 057122 monkeypox v |
| 24 | 176.5 | 21.0 | 349 | 11 | 057109 | 057109 monkeypox v |
| 25 | 176.5 | 21.0 | 349 | 11 | 057110 | 057110 monkeypox v |
| 26 | 176.5 | 21.0 | 349 | 11 | 057111 | 057111 monkeypox v |
| 27 | 176.5 | 21.0 | 349 | 11 | 057112 | 057112 monkeypox v |
| 28 | 176.5 | 21.0 | 349 | 11 | 057113 | 057113 monkeypox v |
| 29 | 176.5 | 21.0 | 349 | 11 | 057114 | 057114 monkeypox v |

| | | | | | | |
|----|-------|------|------|----|--------|--------------------|
| 30 | 176.5 | 21.0 | 326 | 11 | 057122 | 057122 cowpox viru |
| 31 | 176.5 | 21.0 | 351 | 11 | 073559 | 073559 cowpox viru |
| 32 | 176.5 | 21.0 | 349 | 11 | 089118 | 089118 variola vir |
| 33 | 176.5 | 21.0 | 349 | 11 | 089098 | 089098 variola vir |
| 34 | 176.5 | 21.0 | 348 | 11 | 085407 | 085407 variola vir |
| 35 | 175.5 | 20.9 | 349 | 11 | 057305 | 057305 cowpox viru |
| 36 | 174.5 | 20.7 | 349 | 11 | 057284 | 057284 cowpox viru |
| 37 | 174.5 | 20.7 | 349 | 11 | 057097 | 057097 cowpox viru |
| 38 | 174.5 | 20.7 | 349 | 11 | 057098 | 057098 cowpox viru |
| 39 | 173.5 | 20.6 | 351 | 11 | 057117 | 057117 cowpox viru |
| 40 | 167.5 | 19.9 | 350 | 11 | 057123 | 057123 cowpox viru |
| 41 | 144 | 17.1 | 625 | 10 | 035305 | 035305 mus musculu |
| 42 | 128.5 | 15.3 | 267 | 4 | 002764 | 002764 oryctolagus |
| 43 | 124 | 14.7 | 332 | 4 | 077736 | 077736 sus scrofa |
| 44 | 123 | 14.6 | 167 | 11 | 072762 | 072762 rattus norv |
| 45 | 117.5 | 14.0 | 1574 | 10 | 088281 | 088281 rattus norv |

ALIGNMENTS

| RESULT | 1 | PRELIMINARY: | PRT: | 401 AA. |
|-----------------------|---|--------------|------|---------|
| 000300 | 000300 | 000300 | | |
| ID | 000300 | 000300 | | |
| DT | 01-JUL-1997 (TREMBLREL. 04, CREATED) | | | |
| DT | 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE) | | | |
| DT | 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE) | | | |
| DE | OSTEOCALCIN | | | |
| GN | OPG | | | |
| OS | HOMO SAPIENS (HUMAN) | | | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES; | | | |
| OC | CATARRHINI; HOMINIDAE; HOMO. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RX | MEDLINE: 97262071. | | | |
| RA | SIMONET W.S., LACEY D.L., DUNSTAN C.R., KELLEY M., CHANG M.S., | | | |
| RA | LUTHER R., NGUYEN H.Q., WOODEN S., BENNETT L., BOONE T., SHIMAMOTO G., | | | |
| RA | DEROSE M., ELLIOTT R., COLOMBO A., TAN H.L., TRAIL G., SULLIVAN J., | | | |
| RA | DAVEY E., BUCAY N., RENSCHAW-GEGG L., HUGHES T.M., HILL D., PATTERSON W., | | | |
| RA | CAMPBELL P., SANDER S., VAN G., TAPLEY J., DERBY P., LEE R., | | | |
| RA | PROGRAM BOYLE W.J., | | | |
| RT | "Osteoprotegerin: a novel secreted protein involved in the regulation | | | |
| RT | of bone density." | | | |
| RL | CELL 89:309-319(1997). | | | |
| RL | EMBL: U94332; G2072185; . | | | |
| DR | PFAM: PF00020; TNFR_C6; 3. | | | |
| DR | SEQUENCE 401 AA; 45996 MW; 88802A76 CRC32; | | | |
| SO | SEQUENCE | | | |
| Query Match | 36.8%; Score 309.5; DB 2; Length 401; | | | |
| Best local similarity | 41.5%; Pred. No. 4.1e-16; | | | |
| Matches | 56; Conservative 22; Mismatches 52; Indels 5; Gaps 2; | | | |
| QY | 2 ECRRTYVNLGGEREERARACHATHNAPCRRTGFFHAAGFCLEHACPGAGVIAGTGS 61 | | | |
| DB | 78 DEVLXSPVCKELQYKOCNRTNHRVCECKEGRYLEIEFCLEKHRSCPGFVGVOGTGE 137 | | | |
| QY | 62 ONTQOCPCPGPFSSASSSEOCOPRNCALGALNVGSSSHDITLCTGTFPSTRV 121 | | | |
| DB | 138 RNTVCRCPDGPFFNETSKAPCKRHTNCSVGLLTOKGNATHDNI---CSGNSSTOK 194 | | | |
| QY | 122 PGAEE--CERAVIDE 134 | | | |
| DB | 195 CGIDVTLCEAPFRE 209 | | | |
| RESULT | 2 | PRELIMINARY: | PRT: | 401 AA. |
| ID | 060236 | 060236 | | |
| AC | 060236 | 060236 | | |
| DT | 01-AUG-1998 (TREMBLREL. 07, CREATED) | | | |
| DT | 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE) | | | |

[illegible]

```

OC SCURIONGATHI: MORIIDE; MORINAE; MOS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 97262071.
RA SIMONET W.S., LACEY D.L., DUNSTAN C.R., KELLEY M., CHANG M.S.,
RA LUTHEY R., NGUYEN H.O., WOODEN S., BENNETT L., BOONE T., SHIMAMOTO G.,
RA DEPOSE M., ELLIOTT R., COLAMBERO A., TAN H.L., TRAIL G., SULLIVAN J.,
RA DAVE E., BUCK N., RENSHAN-GEIG L., HUGHES T.M., HILL D., PATTISON W.,
RA CAMPBELL P., SANDER S., VAN G., TARELEY J., DERBY P., LEE R.,
RA PROGRAM BOYLE W.J.;
RT "Osteoprotegerin": a novel secreted protein involved in the regulation
RT of bone density.";
RL EMBL; U94331.G2072183; -.
DR MGD; U94331.G2072183; -.
DR PFAM; MG1:109587; OPQ.
DR PFAM; PF00020; TNFR_c6; 3.
SQ SEQUENCE 401 AA; 45923 MW; 8BCBF5B7 CRC32;

Query Match 36.18; Score 303.5; DB 10; Length 401;
Best Local Similarity 38.68; Pred. No. 1.2e-15;
Matches 59; Conservative 24; Mismatches 65; Indels 5; Gaps 2;

QY 2 ERCRCNVLCGEREEDRACHATHNRACRCRTGFFAHNGFGLLEHASCSPRGAVTAPGPS 61
Db 78 DECYCYSVCVELDOSVKOEKRNTHNRVCECEGRILEIEFLKIRHSCPPSGVQAQGP 137
QY 62 QNTQCQPCPCPTGFASSSSSPQCOPHRNCTLGLALNYPGSSSHDTLCSTGFPILSTRV 121
Db 138 RNTVCRKCPDGFSGEITSKAPCIKHTNCSFGLLLIQKGNATHDNV---CSGNREATQK 194
QY 122 PGAE--CERAVIDFVAFQDISTRRLQAL 152
Db 195 CGIDVTLCSEAFAFFRAVPTKIIPNMLSLVDSL 227

```

```

ID 070202 PRELIMINARY; PRT; 401 AA.
AC 070202:
DT 01-AUG-1998 (TREMBLEREL. 07, CREATED)
DT 01-AUG-1998 (TREMBLEREL. 07, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLEREL. 08, LAST ANNOTATION UPDATE)
DE OSTEOCLASTOGENESIS INHIBITORY FACTOR.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA.
OC SCUROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN-NIH SWISS.
RC MIJUNO A., MURAKAMI A., NAKAGAWA N., YASUDA H., TSUDA E., MORINAGA T.,
RA HIGASHIO K.
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-129/OLA;
RX MEDLINE: 98382527.
RA MIZUNO A., MURAKAMI A., NAKAGAWA N., YASUDA H., TSUDA E., MORINAGA T.,
RA HIGASHIO K.;
RT "Structure of the mouse osteoclastogenesis inhibitory factor (OCIF)
gene and its expression in embryogenesis.";
RL GENE 215:339-343(1998).
DR EMBL: AB013898; D1029209; -.
RL EMBL: AB013893; D1034354; -.
DR EMBL: AB013893; D1034354; JOINED.
DR EMBL: AB013900; D1034354; JOINED.
DR EMBL: AB013901; D1034354; JOINED.
DR EMBL: AB013902; D1034354; JOINED.
SQ SEQUENCE 401 AA; 45935 MW; 944D4023 CRC32;

```

```

Query Match 35.7%; Score 300.5; DB 10; Length 401;
Best Local Similarity 37.9%; Pred. No. 1.9e-15;
Matches 58; Conservative 25; Mismatches 65; Indels 5; Gaps 2;

```

```

QY 2 ECRRCNVLCGEREEARACHATHNRACRCRTGFFAHAGFCLLEHASCPPGAGYIAPGIPS 61
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db DECVYCSFVCKELOSKEQCNTHNRVCEGEGRYLEIEFCIKHNSCPGSGSVAGTPE 137
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 62 ONTCCOPCPGPTGFFSASSSSSECCOPHRNCTALGALNVPSSSHDTLCTGCTGFPLSTRY 121
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 138 PNTVCKKCPDGFSESTSKAPCRKHTDCSTFGLLIOKGNATHDNV---CGGNREAIQK 194
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 122 PGAE--CERAVIDEFAFODISIKRLQLLQAL 152
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 195 CGIDVTLCEAFEFRAVPTKILPNMLSVLYDSL 227

```

```

RESULT 6 *
Q16042 PRELIMINARY; PRT; 425 AA.
AC 016042:
DT 01-NOV-1996 (TREMBLEREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLEREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLEREL. 08, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES.
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 91370690.
RA DEMBIC Z., LOETSCHER H., GUBLER U., PAN Y.C., LAHM H.W., GENTZ R.,
RA BROCKHAUS M., LESSLAUER W.;
RT "Two human TNF receptors have similar extracellular, but distinct
intracellular, domain sequences.";
RL CYTOKINE 2:231-237(1990).
DR EMBL: S63368; E91326; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PFAM: PF00020; TNFR_c6; 4.
SQ SEQUENCE 425 AA; 44608 MW; A6B87BF7 CRC32;

```

```

Query Match 28.7%; Score 241; DB 2; Length 425;
Best Local Similarity 36.2%; Pred. No. 5.1e-11;
Matches 46; Conservative 20; Mismatches 49; Indels 12; Gaps 3;

```

```

QY 4 CRYCNVLCGEREEARACHATHNRACRCRTGFFAAG-----FLEHNSCPGAGYIAP 57
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 57 CLSCGSSCDQVETQCTREONRICTPGMYCALSKOEGCRICLAPLKRCPGFGVARP 116
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 58 GTPSONTOCCPPGPTGFSASSSSSECCOPHRNCTALGALNVPSSSHDTLCTGCTGFPL 117
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 117 GRETSDVCKPCAPGPTGFTNTSTSDICRPHQCNVVA---IPGNASMDAVCTST--PT 170
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 118 STRVPGA 124
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 171 RSMAPGA 177

```

```

Query Match 26.3%; Score 221.5; DB 10; Length 459;
Best Local Similarity 37.2%; Pred. No. 1.5e-09;
Matches 42; Conservative 19; Mismatches 41; Indels 11; Gaps 3;

```

```

QY 4 CRYCNVLCGEREEARACHATHNRACRCRTGFF---AHAGF---CLEHASCPPGAGYIA 56
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 79 CLSCSSSCSTQVETQCTREONRICTPGMYCALSKOEGCRICLAPLKRCPGFGVARP 138
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 57 PGTSONTOCCPPGPTGFSASSSSSECCOPHRNCTALGALNVPSSSHDTLCTGCTGFPL 109
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 139 SRAPGNVLCRACAPGPTGFTNTSTSDICRPHQCNVVA---IPGNASMDAVCTST--PT 187

```

```

RESULT 8
O88734 PRELIMINARY; PRT; 482 AA.
AC 088734:
DT 01-NOV-1998 (TREMBLEREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLEREL. 08, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLEREL. 08, LAST ANNOTATION UPDATE)
DE P80 TNF-ALPHA RECEPTOR.
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE: 91370690.
RA DEMBIC Z., LOETSCHER H., GUBLER U., PAN Y.C., LAHM H.W., GENTZ R.,
RA BROCKHAUS M., LESSLAUER W.;
RT "Two human TNF receptors have similar extracellular, but distinct
intracellular, domain sequences.";
RL CYTOKINE 2:231-237(1990).
DR EMBL: S63368; E91326; -.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PFAM: PF00020; TNFR_c6; 4.
SQ SEQUENCE 482 AA; 44608 MW; A6B87BF7 CRC32;

```


Thu May 20 13:49:54 1999

us-09-212-270-2.rspt

Page 6

```
QY      1 LERRCVAVLGGEREERERACHATHNNACCRGFE-----AHAGCLEHASCPPGAGY 54
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      80 LQACLSCGRDSDNOVEFRSNTTHNHCESPEGYLLLGASGCRTCKSKTCKGIYGX 139
        | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY      55 IAPGPSQNTCCOCPCPTGFSSASSSEOCOP--HRNCTALGLANVPSSSHDLCTS 111
        | : | : | | | | | : | : | : | : | : | : | : | : | : |
Db      140 -SGTISGDVYCSCCGPETYSHTVSSDKCEPVVTSTFNTEIVEIML--YPNDTSCTR 196
        | : | : | | | | | : | : | : | : | : | : | : | : | : |
QY      112 CTGFPLSTRVPDAE 125
        | | | : : |
Db      197 TTTTGLSISISTSE 210
```

Search completed: May 14, 1999, 21:39:57
Job time: 7991 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 17:21:17 ; Search time 42.79 Seconds
(without alignments)
99.260 Million cell updates/sec

Title: US-09-212-270-3
Perfect score: 1162
Sequence: 1 MALKVLP.LHRTV.LFAAILFL.....STANGSSRRNP.VAVSHPSV 210

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database: A_Geneseq_34.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|----------------------------|
| 1 | 869 | 74.8 | 417 | 1 | W70386 Amino acid sequenc |
| 2 | 869 | 74.8 | 423 | 1 | W70387 Amino acid sequenc |
| 3 | 155 | 13.3 | 438 | 1 | R81882 Plasmid pDC406/OX4 |
| 4 | 155 | 13.3 | 206 | 1 | R81881 Mouse type-II memb |
| 5 | 155 | 13.3 | 438 | 1 | W48976 OX40/Tc muclein. Pu |
| 6 | 155 | 13.3 | 206 | 1 | W48977 Mouse OX40 extrace |
| 7 | 151 | 13.0 | 176 | 1 | W80254 Amino acid sequenc |
| 8 | 142.5 | 12.3 | 283 | 1 | W05809 Human tumour necro |
| 9 | 142.5 | 12.3 | 283 | 1 | W69238 Hepesvirus entry |
| 10 | 142.5 | 12.2 | 625 | 1 | W69238 Murine NF-kB recep |
| 11 | 142 | 12.2 | 625 | 1 | W68294 Human herpes simpli |
| 12 | 141.5 | 12.2 | 283 | 1 | W12659 Compo virus T2-eq |
| 13 | 141 | 12.1 | 355 | 1 | R85072 Myxoma virus T2 pr |
| 14 | 140 | 12.0 | 326 | 1 | R85072 Human TNF receptor |
| 15 | 139.5 | 12.0 | 283 | 1 | W60045 A receptor protein |
| 16 | 139.5 | 12.0 | 283 | 1 | W79233 NF-kB receptor act |
| 17 | 135 | 11.6 | 616 | 1 | W69952 RANK polypeptide p |
| 18 | 133 | 11.4 | 616 | 1 | W69952 NF-kB receptor act |
| 19 | 133 | 11.4 | 451 | 1 | W69951 RANK polypeptide p |
| 20 | 133 | 11.4 | 451 | 1 | W68287 NF-kB receptor act |
| 21 | 133 | 11.4 | 616 | 1 | W64288 Human DR3 protein. |
| 22 | 130.5 | 11.2 | 833 | 1 | W23220 Extracellular doma |
| 23 | 127.5 | 11.0 | 197 | 1 | R07447 Human laminin B1 c |
| 24 | 126.5 | 10.9 | 466 | 1 | R62655 Compo virus Pst I |
| 25 | 126.5 | 10.9 | 186 | 1 | R07450 Rat Tumour Necrosi |
| 26 | 126 | 10.8 | 461 | 1 | W50893 Human laminin B1 c |
| 27 | 125.5 | 10.8 | 277 | 1 | R38859 CD40 protei |
| 28 | 124.5 | 10.7 | 417 | 1 | W31517 Death domain conta |
| 29 | 121.5 | 10.5 | 428 | 1 | W31517 Human apoptosis pr |
| 30 | 121.5 | 10.5 | 417 | 1 | W31517 Human apoptosis pr |
| 31 | 121.5 | 10.5 | 417 | 1 | W31517 Human apoptosis pr |
| 32 | 121.5 | 10.5 | 417 | 1 | W31517 Human apoptosis pr |
| 33 | 121.5 | 10.5 | 417 | 1 | W31517 Human apoptosis pr |
| 34 | 120.5 | 10.4 | 411 | 1 | W57046 Mouse apoptosis in |
| 35 | 120 | 10.3 | 285 | 1 | W33359 TBP(20-190)/hcg-a1 |
| 36 | 119.5 | 10.3 | 1764 | 1 | P91672 Primary amino acid |
| 37 | 119.5 | 10.3 | 518 | 1 | R51003 Sequence of a reco |
| 38 | 119.5 | 10.3 | 461 | 1 | R51003 p75 Tumour Necrosi |
| 39 | 119.5 | 10.3 | 1776 | 1 | W50894 Mouse laminin B1 c |
| 40 | 118.5 | 10.2 | 461 | 1 | R11141 Human TNF-R deduce |
| 41 | 118.5 | 10.2 | 485 | 1 | R24016 Fusion protein TNF |
| 42 | 118.5 | 10.2 | 461 | 1 | R42058 Filtrblast derived |
| 43 | 117.5 | 10.1 | 256 | 1 | W33357 TBP(20-161)/hcg-a1 |

44 117 10.1 277 1 R76996
45 117 10.1 1798 1 W50896

Deduced sequence e
Human laminin B2 c

ALIGNMENTS

| | |
|--|---|
| RESULT 1 | |
| W70386 | W70386 standard; Protein; 417 AA. |
| AC | W70386: |
| DE | 02-DEC-1998 (first entry) |
| DT | Amino acid sequence of human alpha-OAF065. |
| KW | Human; alpha-OAF065; stroma cell; antibody; inflammatory; |
| OS | Cytokine-mediated disease; rheumatism; ulcerative colitis. |
| OS | Homo sapiens. |
| FH | Key |
| FT | Location/Qualifiers |
| FT | Misc_difference 223 |
| FT | /note= "encoded by NGA" |
| FT | Misc_difference 224 |
| FT | /note= "encoded by CCT" |
| FT | W09838304-A1. |
| PN | 03-SEP-1998. |
| PD | 26-FEB-1998; J00799. |
| PE | 27-FEB-1997; JP-043143. |
| PR | (ONOX) ONO PHARM CO LTD. |
| PI | Fukushima D, Konishi M, Tada H; |
| DR | WPI: 98-481205/41. |
| DR | N-PSDB: V33361. |
| PT | Membrane polypeptide expressed by human stroma cells, and antibodies |
| PT | recognising it - for treatment of inflammatory and other |
| PT | cytokine-mediated diseases. |
| PS | Claim 1; Pages 28-30; 54pp; Japanese. |
| CC | This is the amino acid sequence of the human alpha-OAF065, used in |
| CC | the method of the invention. The process involves the use of peptides |
| CC | expressed by stroma cells, and its antibodies are used for in the |
| CC | prevention and treatment of inflammatory and other cytokine-mediated |
| CC | diseases such as rheumatism, ulcerative colitis. |
| SQ | Sequence 417 AA; |
| Query Match | 74.8%; Score 869; DB 1; Length 417; |
| Best Local Similarity | 83.7%; Pred. No. 6.2e-70; |
| Matches 154; Conservative 11; Mismatches 19; Indels 0; Gaps 0; | |
| QY | 1 MALKVLP.LHRTV.LFAAILFL.LACKVSGTGDCHQOEFKDRSGNCVLCQCGPGWELSK 60 |
| DB | 1 MALKVLP.LHRTV.LFAAILFL.LACKVSGTGDCHQOEFKDRSGNCVLCQCGPGWELSK 60 |
| QY | 61 ECGFGYGEDAQCVPCRPFRFKEDMGFOKPCADCAVNRFORANCSTSDAVCGDCLPG 120 |
| DB | 61 ECGFGYGEDAQCVPCRPFRFKEDMGFOKPCADCAVNRFORANCSTSDAVCGDCLPG 120 |
| QY | 121 FFRKTKLVGFQDMCEVPCGDDPPPEPCHTCKVNLKYSSTVSSPRDRAVAVIGSALAT 180 |
| DB | 121 FFRKTKLVGFQDMCEVPCGDDPPPEPCHTCKVNLKYSSTVSSPRDRAVAVIGSALAT 180 |
| QY | 181 VLLA 184 |
| DB | 181 VLLA 184 |
| RESULT 2 | |
| W70387 | W70387 standard; Protein; 423 AA. |
| AC | W70387: |
| DT | 02-DEC-1998 (first entry) |
| DE | Amino acid sequence of human beta-OAF065. |
| KW | Human; beta-OAF065; stroma cell; antibody; inflammatory; |
| OS | Cytokine-mediated disease; rheumatism; ulcerative colitis. |
| OS | Homo sapiens. |
| FH | Key |
| FT | Location/Qualifiers |
| FT | Misc_difference 223 |

FT Misc-difference /note= "encoded by AGA"
FT 224 /note= "encoded by CCR"
PN MO9838304-A1.
PD 03-SEP-1998.
PF 26-FEB-1998; J00799.
PR 27-FEB-1997; JP-043143.
PA (ONOV) ONO PHARM CO LTD.
PI Fukushima D, Konishi M, Tada H;
DR WPI: 98-481205/41.
N-PSDB: V33362.
PT Membrane polypeptide expressed by human stroma cells, and antibodies
PT recognizing it - for treatment of inflammatory and other
PT cytokine-mediated diseases.
PS Disclosure: Pages 37-49; 54pp; Japanese.
CC This is the amino acid sequence of the human beta-OAF065, used in
CC the method of the invention. The process involves the use of peptides
CC expressed by stroma cells, and its antibodies are used for in the
CC prevention and treatment of inflammatory and other cytokine-mediated
CC diseases such as rheumatism, ulcerative colitis.
SQ Sequence 423 AA;

Query Match 74.8%; Score 869; DB 1; Length 423;
Best Local Similarity 83.7%; Pred. No. 6.3e-70;
Matches 154; Conservative 11; Mismatches 19; Indels 0; Gaps 0;

OY 1 MALKVLPLHRTVLFMAILLHLHACKVSCETGDCRQEFKDRSGNCVLCQGPMEISK 60
ID 1 MALKVLLEDEKFFFLVLGLVYSCKVTEGTGCRQEFKDRSGNCVPCNOCQGPMEISK 60
DB 1 MALKVLLEDEKFFFLVLGLVYSCKVTEGTGCRQEFKDRSGNCVPCNOCQGPMEISK 60
OY 61 ECGFGYGDACVCPRHFRKEDMGFQKCPKADALVNRFORANCSSHTSDAVCGDCLPG 120
ID 61 ECGFGYGDACVCPRHFRKEDMGFQKCPKADALVNRFORANCSSHTSDAVCGDCLPG 120
DB 61 ECGFGYGDACVCPRHFRKEDMGFQKCPKADALVNRFORANCSSHTSDAVCGDCLPG 120
OY 121 FYRKTKLVGFQDMCEVPCGDPPEPYRHCTSKYNLVKISSVSPRDTAAVAAYIC 180
ID 121 FYRKTKLVGFQDMCEVPCGDPPEPYRHCTSKYNLVKISSVSPRDTAAVAAYIC 180
DB 121 FYRKTKLVGFQDMCEVPCGDPPEPYRHCTSKYNLVKISSVSPRDTAAVAAYIC 180
OY 181 VLLA 184
ID 181 VLLA 184
DB 181 VLLA 184

RESULT 3
R81882
ID R81882 standard; Protein: 438 AA.
AC R81882;
DE 30-MAR-1996 (first entry)
DE Plasmid pDC406/OX40/FC* encoding an OX40/FC mutein protein.
KW OX40; OX40-L; cytokine; cell surface molecule; plasmid.
KW pDC406/OX40/FC*; membrane glycoprotein.
OS Synthetic.
PN US5457035-A.
PD 10-OCT-1995.
PF 23-JUL-1993; 097827.
PR 23-JUL-1993; US-097827.
PA (IMMV) IMMUNEX CORP.
PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
DR WPI: 95-357992/46.
N-PSDB: T00829.

PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors
PT and host cells, used to produce recombinant ligand used in e.g.
PT prim. T cell culture, to modulate immune response etc.
PS Example 2: Column 35-38; 26pp; English.
CC This plasmid encodes an OX40/FC antibody fragment mutein protein,
CC and is used to express a soluble OX40/FC mutein fusion protein for
CC use in detecting CDNA clones encoding a OX40 ligand. The FC
CC fragment may be derived from human IgG1, and the plasmid may be
CC used to transform the CV-1/EBNA (ATCC CRL 10478) monkey kidney cell
CC line. Culture supernatant was purified by affinity chromatography
CC and this was used, together with labeled goat anti-human IgG to
CC screen various cell lines.

SQ Sequence 438 AA;

Query Match 13.3%; Score 155; DB 1; Length 438;
Best Local Similarity 27.6%; Pred. No. 1.9e-06;
Matches 59; Conservative 27; Mismatches 88; Indels 40; Gaps 14;

OY 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMEISKEGFGYGDACVPCR 76
ID 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMEISKEGFGYGDACVPCR 76
DB 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMEISKEGFGYGDACVPCR 76
OY 77 PHRFKEDMGFQKCPKADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD 133
ID 77 PHRFKEDMGFQKCPKADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD 133
DB 77 PHRFKEDMGFQKCPKADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD 133
OY 65 TGFYNANVYDICKQCTQCNHSSGSELIKONCTPTDTC-RCRPQTQR-----QDSGY 117
ID 65 TGFYNANVYDICKQCTQCNHSSGSELIKONCTPTDTC-RCRPQTQR-----QDSGY 117
DB 65 TGFYNANVYDICKQCTQCNHSSGSELIKONCTPTDTC-RCRPQTQR-----QDSGY 117
OY 133 ---MECVPCGDPPEPYEP---HCTSKYNLVKISSVSPRDTAAVAAYIC---SALATVL 182
ID 133 ---MECVPCGDPPEPYEP---HCTSKYNLVKISSVSPRDTAAVAAYIC---SALATVL 182
DB 133 ---MECVPCGDPPEPYEP---HCTSKYNLVKISSVSPRDTAAVAAYIC---SALATVL 182
OY 118 KLGVCVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
ID 118 KLGVCVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
DB 118 KLGVCVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
OY 183 LACS-----SCVSSFARSSSRRRPAAVSSHPV 210
ID 183 LACS-----SCVSSFARSSSRRRPAAVSSHPV 210
DB 183 LACS-----SCVSSFARSSSRRRPAAVSSHPV 210
OY 175 WETQRPFRPTVOST--TWMPRTSELPSTPTL 205
ID 175 WETQRPFRPTVOST--TWMPRTSELPSTPTL 205
DB 175 WETQRPFRPTVOST--TWMPRTSELPSTPTL 205

RESULT 4
R81881
ID R81881 standard; Protein: 206 AA.
AC R81881;
DE 08-JUL-1996 (first entry)
DE Mouse type-II membrane polypeptide OX40 extracellular domain.
KW OX40; OX40-L; cytokine; cell surface molecule;
KW membrane glycoprotein.
OS Mus musculus.
PN US5457035-A.
PD 10-OCT-1995.
PF 23-JUL-1993; 097827.
PR 23-JUL-1993; US-097827.
PA (IMMV) IMMUNEX CORP.
PI Baum PR, Fanslow WC, Gayle RB, Goodwin RG;
DR WPI: 95-357992/46.
N-PSDB: T00826.

PT New isolated DNA encoding the OX40 ligand polypeptide - also vectors
PT and host cells, used to produce recombinant ligand used in e.g.
PT prim. T cell culture, to modulate immune response etc.
PS Example 1: Column 33-34; 26pp; English.
CC This sequence encodes the extracellular domain of OX40, a membrane
CC glycoprotein present on the CD4 positive subset of activated T
CC cells.
SQ Sequence 206 AA;

Query Match 13.3%; Score 155; DB 1; Length 206;
Best Local Similarity 27.6%; Pred. No. 8.1e-07;
Matches 59; Conservative 27; Mismatches 88; Indels 40; Gaps 14;

OY 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMEISKEGFGYGDACVPCR 76
ID 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMEISKEGFGYGDACVPCR 76
DB 18 LFLHLHACKVSCETGDCRQEFKDRSGN-CVLCQKQGPMEISKEGFGYGDACVPCR 76
OY 77 PHRFKEDMGFQKCPKADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD 133
ID 77 PHRFKEDMGFQKCPKADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD 133
DB 77 PHRFKEDMGFQKCPKADALVNRFO-RANCSHTSDAVCGDCLPGFYRKTKLVGFOD 133
OY 65 TGFYNANVYDICKQCTQCNHSSGSELIKONCTPTDTC-RCRPQTQR-----QDSGY 117
ID 65 TGFYNANVYDICKQCTQCNHSSGSELIKONCTPTDTC-RCRPQTQR-----QDSGY 117
DB 65 TGFYNANVYDICKQCTQCNHSSGSELIKONCTPTDTC-RCRPQTQR-----QDSGY 117
OY 133 ---MECVPCGDPPEPYEP---HCTSKYNLVKISSVSPRDTAAVAAYIC---SALATVL 182
ID 133 ---MECVPCGDPPEPYEP---HCTSKYNLVKISSVSPRDTAAVAAYIC---SALATVL 182
DB 133 ---MECVPCGDPPEPYEP---HCTSKYNLVKISSVSPRDTAAVAAYIC---SALATVL 182
OY 118 KLGVCVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
ID 118 KLGVCVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
DB 118 KLGVCVPC--PPGHSPGNNQACKPWTNCTLSGKQTRHPADSIDAV-CEDRSLATL 174
OY 183 LACS-----SCVSSFARSSSRRRPAAVSSHPV 210
ID 183 LACS-----SCVSSFARSSSRRRPAAVSSHPV 210
DB 183 LACS-----SCVSSFARSSSRRRPAAVSSHPV 210
OY 175 WETQRPFRPTVOST--TWMPRTSELPSTPTL 205
ID 175 WETQRPFRPTVOST--TWMPRTSELPSTPTL 205
DB 175 WETQRPFRPTVOST--TWMPRTSELPSTPTL 205

RESULT 5

PT for treatment of bone growth disorders
 PS Claim 1; Pages 29-31; 51pp; Japanese.
 CC The present sequence represents a protein designated 7p4. This protein
 CC is capable of inducing differentiation in osteoblast cells. The
 CC protein may be used to screen compounds for the ability to bind to
 CC it, for use as ligands, agonists or antagonists and inhibiting or
 CC otherwise altering its differentiation inducing activity. Compounds
 CC so identified, as well as the protein itself, DNA encoding it, and
 CC antibodies to it, may be used in the treatment of diseases of bone
 CC growth and osteoblast differentiation, such as bone sarcomas.
 SQ Sequence 176 AA;

Query Match 13.0%; Score 151; DB 1; Length 176;
 Best Local Similarity 29.9%; Pred. No. 1,5e-06;
 Matches 49; Conservative 22; Mismatches 63; Indels 30; Gaps 10;

QY 13 LFAILFLHLACKVSCETGDCRQOEFKDRSGNCVLCRCGPGMELSKGCGFYGEDAQC 72
 || :
 Db 22 LFLVIRAMPESYSFNCPPDEYQ-----SNDVCKCTGSGTFVAPCKIPIHTQ-GQC 72
 QY 73 VPCPHPEFK-EDMGFOCKPCADALVNRPR--ANGSHSDAVCGDCLPEFYRKTALVG 129
 || :
 Db 73 EKCHPGFTFGKNDLHDCLELCTC--DKDGMVADCSATSDRC-RCQIGLY----- 122
 QY 130 FQDME---CVPCGDP--PYEPCTSKVLYKISVSSPRD 167
 || :
 Db 122 YDPRKFPESCRCPTKCPQGITVLOECNSTANTV-CSSSVSNPRN 164

RESULT 8
 ID W05809 standard; Protein; 283 AA.
 AC W05809;
 DT 30-JUN-1997 (first entry)
 DE Human tumour necrosis factor receptor.
 KW Human; tumour necrosis factor; TNF; receptor; treatment;
 KW activation; inhibition; identification; agonist; antagonist;
 KW stimulation; T cell; differentiation; mediation; immune;
 KW antiviral; response; regulation; growth; protection; HIV;
 KW radiation; Chlamydia; infection; immunodeficiency; virus;
 KW autoimmune disease; inflammation; septic shock; cerebral; malaria;
 KW cachexia; B cell; cancer; graft; host; reaction; rejection;
 KW prevention; apoptosis; cytotoxicity; hybridisation probe;
 KW detection; antibody; reagent.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT peptide 1..38
 FT /label= sig_peptide
 FT 39..283
 FT /label= mat_peptide
 PN W09634095-A1.
 PD 31-OCT-1996.
 PE 27-APR-1995; U05058.
 PR 27-APR-1995; WO-U05058.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Gentz R, Ni J, Rosen CA,
 DR WPI; 96-497627/49.
 DR N-PSDB; T40048.
 PT New nucleic acid encoding a human tumour necrosis factor receptor -
 PT useful for treatment of auto-immune diseases etc., in diagnosis and
 PT for drug screening
 PS Claim 1; Pages 47-48; 73pp; English.
 CC The present sequence is the human tumour necrosis factor (TNF)
 CC receptor, which may be used for the identification of TNF receptor
 CC agonists or antagonists. TNF receptor agonists inhibit tumour
 CC growth, stimulate cell (e.g. T cell) differentiation, mediate the
 CC immune and antiviral responses, regulate growth, protect against
 CC radiation and Chlamydia infection, and can be used to treat
 CC immunodeficiencies, e.g. human immunodeficiency virus. Antagonists
 CC can be used to treat T cell mediated autoimmune disease,
 CC inflammation, septic shock, cerebral malaria, cachexia or B cell
 CC cancers, to inhibit graft-host reactions and to prevent apoptosis

CC or cytotoxicity. Fragments of the TNF receptor encoding DNA can be
 CC used as hybridisation probes for detecting related genes.
 CC Antibodies against the protein can be used as reagents for
 CC detecting/measuring soluble forms of protein in the circulation.
 SQ Sequence 283 AA;

Query Match 12.3%; Score 142.5; DB 1; Length 283;
 Best Local Similarity 25.4%; Pred. No. 1,5e-05;
 Matches 60; Conservative 28; Mismatches 83; Indels 65; Gaps 16;

QY 10 RTVLAAILFLHLACKVSCET-GDCRQOEFKDRSGNCVLCRCGPGMELSKGCGFYGE 68
 || :
 Db 17 RTDVLRLVILTLTGAPCAVPAALPSCKEDYPPGS-EC--CPKCSGFRYKKEKCGELTG- 73
 QY 69 DAOCVPCPRPRFREDW-GFOCKRPCADC--ALVNRFORANGSHTSVAVCGDCLPGFY-- 123
 || :
 Db 73 -TVCEPCPPETYLAIHLNGLSKLCQCCQCDPAMGLRASR-NCSTENAVCG-CSPGHICIV 129
 QY 123 -----RRTKLVG--FQDMECVPCGDPPEPIEPHCT-----SKVN- 155
 || :
 Db 130 QDSDHCAACRAVATSSPGQVQKGTESDPTLQNC--PPGTFSPNGTLECOHQTKCSW 187
 QY 155 -LVKISVSSPR-----DFAVAVIGSALATVLIACSSCVSSTARSSWRNP 202
 || :
 Db 188 LVTKAGAGTSSHWMMFLSGSLVYIVCVSTVGLIT-----CVK-----RRKP 230

RESULT 9
 ID W69238 standard; Protein; 283 AA.
 AC W69238;
 DT 21-OCT-1998 (first entry)
 DE Herpesvirus entry mediator protein.
 KW Entry mediator gene; herpesvirus; HVEM; tumour necrosis factor receptor;
 KW gene expression regulator; cellular stress; inflammatory response;
 KW lymphocyte activity regulator; autoimmune response.
 OS Homo sapiens.
 PN W09825967-A1.
 PD 18-JUN-1998.
 PE 05-DEC-1997; U22278.
 PR 12-DEC-1996; US-032705.
 PA (GETH) GENENTECH INC.
 PI Ashkenazi AJ, Marsters SA,
 DR WPI; 98-348457/30.
 DR N-PSDB; V44852.
 PT Herpesvirus entry mediator polypeptide, HVEM - useful, e.g. in
 PT assays for HVEM and to produce antibodies and transgenic animals,
 PT e.g. for drug screening
 PS Claim 1; Fig 1; 46pp; English.
 CC This sequence is the herpesvirus entry mediator (HVEM) protein of
 CC the invention. The protein is useful in quantitative diagnostic assays
 CC for HVEM, in affinity purification of HVEM from recombinant cells/natural
 CC sources and in competitive-type receptor binding assays. It can also be
 CC used to generate antibodies, also useful in diagnostic assays for HVEM
 CC and affinity purification of HVEM. HVEM is believed to be a member of the
 CC tumour necrosis factor receptor (TNFR) family, and transient transfection
 CC of HVEM into human 293 cells caused marked activation of certain
 CC transcription factors, e.g. AP-1, suggesting that HVEM is involved in
 CC regulating gene expression in response to infectious stimuli and cellular
 CC stress. The predominant expression of HVEM mRNA in lymphocyte-rich
 CC tissues (e.g. spleen and peripheral blood) also suggested it may be a
 CC receptor in regulating lymphocyte activity. Antibodies produced may be
 CC therefore be useful therapeutically, e.g. antagonistic antibodies may be
 CC useful to block excessive inflammatory/autoimmune response resulting from
 CC e.g. AP-1 induction, whilst agonistic antibodies may enhance HVEM
 CC regulation of such induction. The DNA may be used diagnostically, e.g.
 CC to determine if DNA and/or RNA encoding HVEM is present in cells, and to
 CC prepare HVEM polypeptide recombinantly. It is also useful to produce
 CC non-human transgenic animals (e.g. mice or rats), especially knockout
 CC animals containing cells with an altered gene encoding HVEM polypeptide.
 CC Such animals are useful in the development and screening of
 CC therapeutically useful reagents.

| | | |
|-----------------------|---|--|
| SQ | Sequence | 283 AA: |
| Query Match | 12.3%; | Score 142.5; DB 1; Length 283; |
| Best Local Similarity | 25.4%; | Pred. No. 1.5e-05; |
| Matches | 60; Conservative | 28; Mismatches 83; Indels 65; Gaps 16; |
| OY | 10 RTVFAALIFLHLACKVSCET-GDCRQGEFKDRSGNCVLKQCQGFMELSKCEGFGYE | 68 |
| Dd | 17 RTDVRRLVLYFLFLGAPCYAPALPSCKEDEYPVG--EC--CPKCSGYRVKACGLTG- | 73 |
| OY | 69 DAQCPCRHRRKEEM-GRQCKPCADC--ALVNFRQANCSTHSAYVCDDLPEFY---- | 123 |
| Dd | 73 -TVCEPCPGPTIYAHNLNGSLKCLOCMCPAMGLRASR-NCSRTEAAVCG-CSPGFCTIV | 129 |
| OY | 123 -----RKTLVG--FDDMCVPCGDPPPYEPHCT-----SKVN- | 155 |
| Dd | 130 QDGDIACACRAATATSSPGCRVOKGGTESDITLCQNC--PPGFPSPGTIEEOCHOTKCSW | 187 |
| OY | 155 -LVKISIVSVSR-----DTAAVAITCASALATVLLACCSCVSSTARGSWMRRNP | 202 |
| Dd | 188 LVTKAGAGTSSHHWWMFLSGLVIYICSTGLTI-----CVK-----RRKP | 230 |
| RESULT | 10 | |
| ID | M69958 | |
| AC | M69958 standard; Protein: 625 AA. | |
| DT | 08-OCT-1998 | (first entry) |
| DE | Murine NF-kB receptor activator mURANK. | |
| KM | RANK; necrosis factor kappA B; NF-kB; receptor activator; human; | |
| KW | immune response; inflammatory response; toxic shock; sepsis; | |
| RN | RANKL; RANK ligand; tumour necrosis factor; TNF; murine. | |
| OS | Mus sp. | |
| PN | WO9828426-A2. | |
| PD | 02-JUL-1998 | |
| PF | 22-DEC-1997; U23775. | |
| PR | 14-OCT-1997; US-064671. | |
| PR | 23-DEC-1996; US-059978. | |
| PA | 07-MAR-1997; US-813509. | |
| PI | (IMMV) IMMUNEX CORP. | |
| DR | Anderson DM, Galibert LJ, Maraskovsky E; | |
| DR | WP1: 96-377657/32. | |
| DR | N-PDSB: V41379. | |
| PT | New isolated ligand for receptor activator of NF-kappa B - used to | |
| PT | develop products for augmenting an immune response for inhibiting an | |
| PT | inflammatory response and for protection of cells | |
| BS | Example 14; Pages 62-64; 80pp; English. | |
| CC | This represents the murine mURANK (receptor activator of necrosis | |
| CC | factor-kappaB (NF-kB)) polypeptide which is a homolog of the human RANK. | |
| CC | RANK is a member of the tumour necrosis factor (TNF) family. A soluble | |
| CC | RANK may be used for inhibiting activation of NF-kB, by contacting a cell | |
| CC | expressing membrane-associated RANK with a soluble RANK which binds to | |
| CC | RANK ligand (RANKL). RANKL polypeptides can activate RANK and can be | |
| CC | used to induce maturation of dendritic cells and enhance their | |
| CC | allo-stimulatory capacity, thereby augmenting an immune response. The | |
| CC | soluble RANK polypeptide composition may also be used for regulating an | |
| CC | immune or inflammatory response. Inhibition of NF-kB by RANK antagonists | |
| CC | may be useful in ameliorating negative effects of an inflammatory | |
| CC | response that result from triggering of RANK, e.g. in treating toxic | |
| CC | shock or sepsis, graft-versus-host reactions, or acute inflammatory | |
| CC | reactions. They can also be used in adjunct therapy for disease | |
| CC | characterised by neoplastic cells that express RANK. RANKL polypeptides | |
| CC | can also be used to identify inhibitors of RANK and thus inhibitors of | |
| CC | an inflammatory response, and also for protecting RANK-expressing cells | |
| CC | from the negative effects of chemotherapy and the presence of high levels | |
| CC | of TNF-alpha. The products can also be used for detection and drug | |
| CC | screening. | |
| SQ | Sequence | 625 AA: |

| | | | | | | | | | |
|-----------------------|---|--|------------|------------|-----|--------|-----|------|-----|
| Matches | 54; | Conservative | 26; | Mismatches | 76; | Indels | 44; | Gaps | 12; |
| OY | 13 | LFAALLFLHLHACKVSCSTGDCRQEFKDRSGNCVLAKCQSPGMELSKEGCFGYGEDAQC | 72 | | | | | | |
| Dd | 15 | LLALCVLLVPLIQVTLQV--TPPCTOGRNHEHLGRC--CSRCEBKGKLSSKC--TPTSDVC | 69 | | | | | | |
| OY | 73 | VPCRPFRHKEDWGFQ-KC---KPC-ADCALV-----NRFQRANCSHT-----SDAVC-- | 115 | | | | | | |
| Dd | 70 | LPCGGDEYLDITWNNEEDCKLHKVCDAGAKALVAVDPGNHTAPRRCACTAGYHMNSDCECCR | 129 | | | | | | |
| OY | 115 | --GDGLPEFYRKTKLYGVQDMECVYC-----GDPPFPYPHPCTISKNVNLKIST | 161 | | | | | | |
| Dd | 130 | RNTFCAPFPGAOHPQLQNKDYTCPTCLLGFSDFVSFTDKCPW-TNCT-----LLGK | 181 | | | | | | |
| OY | 162 | VSSPRDTAAVAVICSAATV | 181 | | | | | | |
| Dd | 182 | LEAHGTESDVYCGSSMTL | 201 | | | | | | |
| RESULT | 11 | | | | | | | | |
| W68294 | | | | | | | | | |
| ID | W68294 | standard; protein; 625 AA. | | | | | | | |
| AC | W68294: | | | | | | | | |
| DI | 08-OCT-1998 | (first entry) | | | | | | | |
| KW | Murine NF-kB receptor activator mURANK. | | | | | | | | |
| RV | RANK; necrosis factor-kappa B; NF-kB; receptor activator; human; | | | | | | | | |
| OS | immune response; inflammatory response; toxic shock; sepsis; | | | | | | | | |
| KS | RANKL; RANK ligand; tumour necrosis factor; TNF; murine. | | | | | | | | |
| SW | Mus SP. | | | | | | | | |
| PN | WO9828424-A2. | | | | | | | | |
| PD | 02-JUL-1998. | | | | | | | | |
| PF | 22-DEC-1997; U23866. | | | | | | | | |
| PR | 14-OCT-1997; US-064671. | | | | | | | | |
| PR | 23-DEC-1996; US-059878. | | | | | | | | |
| PR | 07-MAR-1997; US-813509. | | | | | | | | |
| PA | (IMMUNEX CORP. | | | | | | | | |
| DR | Anderson DM, Galibert LJ, Maraskovsky E; | | | | | | | | |
| DR | N-PSDB: VAI373. | | | | | | | | |
| PI | New isolated receptor activator of necrosis factor-kappa B - useful | | | | | | | | |
| PT | for, e.g. developing products for regulating an immune or | | | | | | | | |
| P1 | inflammatory response, treating toxic shock or sepsis | | | | | | | | |
| PS | Claim 27; Pages 62-64; 80pp; English. | | | | | | | | |
| CC | This represents the murine muRANK (receptor activator of necrosis | | | | | | | | |
| CC | factor-kappab (NF-kB)) polypeptide which is a homolog of the human RANK. | | | | | | | | |
| CC | RANK is a member of the tumour necrosis factor (TNF) family. Host cells | | | | | | | | |
| CC | transformed or transfected with an expression vector comprising the RANK | | | | | | | | |
| CC | encoding nucleic acid can be used to produce recombinant RANK protein. | | | | | | | | |
| CC | The soluble RANK may be used for inhibiting activation of NF-kB, by | | | | | | | | |
| CC | contacting a cell expressing membrane-associated RANK with a soluble | | | | | | | | |
| CC | RANK which binds to RANK ligand (RANKL). The soluble RANK polypeptide | | | | | | | | |
| CC | composition may also be used for regulating an immune or inflammatory | | | | | | | | |
| CC | response. Inhibition of NF-kB by RANK antagonists may be useful in | | | | | | | | |
| CC | ameliorating negative effects of an inflammatory response that result | | | | | | | | |
| CC | from triggering of RANK, e.g. in treating toxic shock or sepsis, graft- | | | | | | | | |
| CC | -versus-host reactions, or acute inflammatory reactions. They can also be | | | | | | | | |
| CC | used in adjunct therapy for disease characterised by neoplastic cells | | | | | | | | |
| CC | that express RANK. The products can also be used for detection and drug | | | | | | | | |
| CC | screening. | | | | | | | | |
| SQ | Sequence | 625 AA; | | | | | | | |
| Query Match | | 12.2%: Score 142; DB 1; Length 625; | | | | | | | |
| Best Local Similarity | 27.0%: Pred. No. 4e+05; | | | | | | | | |
| Matches | 54; Conservative | 26; Mismatches | 76; Indels | 44; Gaps | 12; | | | | |
| OY | 13 | LFAALLFLHLHACKVSCSTGDCRQEFKDRSGNCVLAKCQSPGMELSKEGCFGYGEDAQC | 72 | | | | | | |
| Dd | 15 | LLALCVLLVPLIQVTLQV--TPPCTOGRNHEHLGRC--CSRCEBKGKLSSKC--TPTSDVC | 69 | | | | | | |
| OY | 73 | VPCRPFRHKEDWGFQ-KC---KPC-ADCALV-----NRFQRANCSHT-----SDAVC-- | 115 | | | | | | |
| Dd | 70 | LPCGGDEYLDITWNNEEDCKLHKVCDAGAKALVAVDPGNHTAPRRCACTAGYHMNSDCECCR | 129 | | | | | | |
| OY | 115 | --GDGLPEFYRKTKLYGVQDMECVYC-----GDPPFPYPHPCTISKNVNLKIST | 161 | | | | | | |
| Dd | 130 | RNTFCAPFPGAOHPQLQNKDYTCPTCLLGFSDFVSFTDKCPW-TNCT-----LLGK | 181 | | | | | | |
| OY | 162 | VSSPRDTAAVAVICSAATV | 181 | | | | | | |
| Dd | 182 | LEAHGTESDVYCGSSMTL | 201 | | | | | | |
| OY | 73 | VPCRPFRHKEDWGFQ | | | | | | | |


```

CC      to its receptor.  T2 protein is useful (when administered by
CC      injection or infusion from sustained release implants, etc.) for
CC      treating TNF associated toxicity e.g. side effects of using TNF as
CC      an antitumour agent or in cases of cachexia and septic shock where
CC      TNF production is excessive.
CC      Sequence 326 AA;

OY      13 LFAALLFLHLAC-----KVCSEGDGRQDFEKRSGNVLCKKCGGMEISKCCGFGY 66
Dd      1 MERLLTLAAVYACVYGAGPYGADRGKCRGNDY-EKDSLC-CTSCPPGSYASRLC--GP 55

OY      67 GEDACVCPCHRFREDFMGF-QKRCPCADCALVYNFQANCSHTSDAIVCGDLPBFY--- 123
Dd      56 GSDIVCSFCKMETFLASTINHPACVSCGRGCTGHLSEOSCDKTRDRC-DCSAGNYCLL 114

OY      123 -----RKTKL-VGF-----QDMCEVPCGDDPPPY-----BPHGTSKVL 155
Dd      115 KGECCRICAPKTKCPAGIGVSGHTRIGDVLCTKC---PRITYSDAVSTETCTSFENY 170

OY      156 VKISSTVSSPRTIAVAAVICSALA 179

```

Search completed: May 13, 1999, 17:21:18
Job time: 20717 sec

This Page Blank (uspto)

| | | | | | |
|----|-----|--|--|-------------|-----|
| QY | 111 | DAVCG----- | -DCLPGFYRKTKL----- | -VGF--QMECV | 137 |
| Db | 122 | DYVCGCKENQFORYLSETHFOCVDCSPCFNGTIVTIPCKETQNTVNCNCHAGFLRESECV | 181 | | |
| QY | 138 | CGDPPPEPEPHCTSVNLYKTS----- | STVSSPRTAAVAAYICSLAIVLLACSSC-VSS | 191 | |
| Db | 182 | CS----- | HCKKNECKMLCLPPLPLANVNPQDSGAVLLPLVILLGLCLLSEIFISL | 233 | |
| QY | 192 | TARGSSWRNPAYS | 206 | | |
| Db | 234 | MCRYPRMR--PEVTS | 246 | | |

RESULT 7
D36858
gene GAr protein - variola virus
N:Alternate names: B28R protein (COP)
C:Species: variola virus
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 10-Sep-1997
C:Accession: D36858; S46888; S32385; S35987
R:Blinov, V.M.
submitted to GenBank, November 1992
A:Description: not shown.
A:Reference number: A36859
A:Accession: D36858
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <BLI>
A:Cross-references: GB:X69198; NID:q456758; PID:q457087
A:Experimental source: strain India-1967, ssp. major, isolate Ind3
R:Kolychalov, A.A.; Blinov, V.M.; Gytarov, V.V.; Pozdnyakov, S.G.; Chizhikov, V.E.; F
Submitted to the EMBL Data Library, April 1992
A:Description: Nucleotide sequence analysis of the region of variola virus XhoI F O H
A:Reference number: S46888
A:Accession: S46888
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-349 <KOL>
A:Cross-references: EMBL:X67117; NID:q516428; PID:q516449
A:Experimental source: strain India-1967, isolate Ind3
R:Shchelkunov, S.N.; Blinov, V.M.; Sandakhchiev, L.S.
FEBS Lett. 319, 80-83, 1993
A:Title: Genes of variola and vaccinia viruses necessary to overcome the host protective
A:Reference number: S32385
A:Accession: S32385
A:Molecule type: DNA
A:Residues: 31-168 <SHC>
A:Cross-references: EMBL:X69198
A:Experimental source: strain India-1967, ssp. major
C:Genetics:
A:Gene: GAr
C:Superfamily: NGF receptor repeat homology
F:68-109/Domain: NGF receptor repeat homology <NG2>
F:110-151/Domain: NGF receptor repeat homology <NG3>

Query Match 11.9%; Score 138.5; DB 2; Length 349;
 Best Local Similarity 23.7%; Pred. No. 0.00011;
 Matches 50; Conservative 30; Mismatches 88; Indels 43; Gaps 10;

QY 10 RTVEFAILLFLHL-----ACKVSCETGDCROQFEDRSNGCVLCKOCGPMELSKREG 63
 Db 2 KSVLLYLTLFLSCILINRGDAPYTPPPKCKDTEYK-RHNLIC--CLSCPGTYASRLC- 58
 QY 64 FGYGDAOCVPCRRHRF-KEDMGFOKCKPCADCALVNRFORANCSTSDAVCGDCLPGFY 122
 Db 58 -DSKTNCTCTCGSGTFTSRNNHLPACLSGRCNSNVETRSCTTHNRIC-ECSPGY 115
 QY 123 -----RRTKL-VGF-----QMECVPCG-----DPPPEPHCTSK 152
 Db 116 CLKSSGCKACVSGTKGIGGVSGHVSQVIGTSFGFGYTSHTVSADKCEVPVNNVT 175
 QY 153 VNLKISSTVSSPRDTAAVAVICSAATATVL 183
 Db 176 FNYIDVEITLVPNDTCTRTTGTLSL 206

RESULT 8

Query Match 11.4%; Score 132; DB 2; Length 416;
 Best Local Similarity 27.8%; Pred. No. 0.00043;
 Matches 59; Conservative 19; Mismatches 100; Indels 34; Gaps 11;

QY 21 LHLACKVSCETGD---CROGEFKDR-SGNCVLCCKCGPMELSKREGCGY-----GED 69
 Db 84 LHSMAKPCVSDAYCAKAYGFDLSSGCKECSIC-----EVGGLMFPDRSDOD 135
 QY 70 AOCVPCRRHRKEDMGF-QCKPCADCALVNRFORANCSTSDAVCGDCLPGFYKRT-KL 127

A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 21-35,'Y','R',174-275,'S','277-395,'R',397-416 <HEU>
 C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma
 C:Comment: The cysteine-rich region of the extracellular domain may form part or all of
 C:Comment: This protein is thought to form a high-affinity receptor when it associates w
 C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor;
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:21-416/Product: nerve growth factor receptor #status predicted <MAT>
 F:21-239/Domain: extracellular #status predicted <EXT>
 F:24-57/Domain: NGF receptor repeat homology <NG1>
 F:59-100/Domain: NGF receptor repeat homology <NG2>
 F:101-139/Domain: NGF receptor repeat homology <NG3>
 F:141-181/Domain: NGF receptor repeat homology <NG4>
 F:189-237/Region: serine/threonine-rich
 F:240-261/Domain: transmembrane #status predicted <MEM>
 F:262-416/Domain: intracellular #status predicted <INT>
 F:52/Binding site: carbohydrate (Asn) (covalent) #status predicted

Db 136 TVCEECPEGTSDANFVDPCLPCTICE-ENEMVWKECTATSDAECRDLHPRTHTHPSL 194
 QY 128 VGFQDMCEVPCGDDPPPEPHCTSKVNLVKSIVSSP---RDR--VAAVICSAATATVL 182
 Db 195 AGSDSP--PITRDFEFTGNATTLADITVTVMGSSOPVSRGTADNLIPIYCSILAAYV 252
 QY 183 IACSSCV-----STRAGSSWRNDPAVSSHP 209
 Db 253 VGLAVYIAFKRMNSCKOKKCANRPNVNOTPS 284

RESULT 9

Query Match 11.1%; Score 129; DB 1; Length 461;
 Best Local Similarity 22.1%; Pred. No. 0.00086;
 Matches 55; Conservative 31; Mismatches 109; Indels 54; Gaps 10;

QY 6 LPHRTVFAILLFLHLACKVSCETG-----DCROEFKDRSGNCVLCCKOC 52
 Db 3 LPIVGLLSIVLLALLMGHPSGVTGLVPSLGRKRDNLCPGKVAHPKNNISICTKC 62
 QY 53 GPMELSKREGCGFGEDACVPCRRHRKEDMG-FQCKPCADCAL-VNFORANCSTSD 110
 Db 63 HKGYLVSDCP-SPOGEVCEVCCKRGFTTASQNHVROCLCKTRKEMFVEISPCRAM 121
 QY 111 DAVCGDCLPGFYRTKLVGFQDMCEVPC-----GDPEPP 144
 Db 122 DTVCGCKKNQORILSETHFCVDCSPFCNCTVITPCKEKONTVCNCHAGFLLSGNECTP 181
 QY 145 YEPHCTSKVNLVKS---STVSSPRDTAAVAVICSAATATVLAC---SSCVSSTARSS 197
 Db 182 CS-HCKKNQECMKLCLPVAVNTNPDSGTAVLL-PLVIFLGCLLFFLICISLGRPQ 238
 QY 198 WRNDPAVSS 206
 Db 239 WR--PRVYS 245

A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-461 <HTM>
 A:Cross-references: GB:M63122; NID:g207361; PID:g207362
 C:Comment: This protein is one of two known receptors for both TNF-alpha (cachectin)
 C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
 C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-461/Product: tumor necrosis factor receptor type 1 #status predicted <MAT>
 F:30-201/Product: tumor necrosis factor binding protein #status predicted <TBP>
 F:44-82/Domain: NGF receptor repeat homology <NG1>
 F:84-126/Domain: NGF receptor repeat homology <NG2>
 F:127-167/Domain: NGF receptor repeat homology <NG3>
 F:168-204/Domain: NGF receptor repeat homology <NG4>
 F:212-234/Domain: transmembrane #status predicted <MEM>
 F:235-461/Domain: intracellular #status predicted <INT>
 F:54,151,201/Binding site: carbohydrate (Asn) (covalent) #status predicted

RESULT 10

Query Match 11.1%; Score 129; DB 1; Length 461;
 Best Local Similarity 22.1%; Pred. No. 0.00086;
 Matches 55; Conservative 31; Mismatches 109; Indels 54; Gaps 10;

A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-461 <HTM>
 A:Cross-references: GB:M63122; NID:g207361; PID:g207362
 C:Comment: This protein is one of two known receptors for both TNF-alpha (cachectin)
 C:Superfamily: tumor necrosis factor receptor type 1; NGF receptor repeat homology
 C:Keywords: duplication; glycoprotein; receptor; transmembrane protein
 F:1-29/Domain: signal sequence #status predicted <SIG>
 F:30-461/Product: tumor necrosis factor receptor type 1 #status predicted <MAT>
 F:30-201/Product: tumor necrosis factor binding protein #status predicted <TBP>
 F:44-82/Domain: NGF receptor repeat homology <NG1>
 F:84-126/Domain: NGF receptor repeat homology <NG2>
 F:127-167/Domain: NGF receptor repeat homology <NG3>
 F:168-204/Domain: NGF receptor repeat homology <NG4>
 F:212-234/Domain: transmembrane #status predicted <MEM>
 F:235-461/Domain: intracellular #status predicted <INT>
 F:54,151,201/Binding site: carbohydrate (Asn) (covalent) #status predicted

| | | | | |
|--------------------------|-------|-------------------|-----------|------------|
| Query Match | 11.1% | Score 128.5 | DB 2 | Length 277 |
| Best Local Similarity | 24.9% | Pred. No. 0.00061 | | |
| Matches 54, Conservative | 30 | Mismatches 82 | Indels 51 | Gaps 11 |

RESULT 11

F:171-179/Region: proline/serine/threonine-rich
F:183-202/Domain: transmembrane #status predicted <TM>
F:203-250/Domain: intracellular #status predicted <INT>
F:95,162/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.8%; Score 125.5; DB 1; Length 250;
Best Local Similarity 26.9%; Pred. No. 0.00099;
Matches 49; Conservative 21; Mismatches 59; Indels 53; Gaps 12;

OY 34 CRGQEFKDRSGNCVLCRGGCGMELSKRCGCYGED---AACVCORP-HRKEDMGFG-K 88
Db 27 CPDHHWTGGGLC--CRRCCEPGTFEVKDC---BODRTAAQCDDPCIPETSSPDYHTRPH 80

OY 89 CKPADCALVRFORFANCSHTSDAVCGCDLGPFRKTLYEFOFMGVPCG----- 140
Db 81 CESGRHCN-SGFILRNCTVTANAEC-SC-----SKNMCCRGDETCDDPLPALTR 130

OY 140 -----DPPPEYPCHCKT---VLVKISSLVSSEPRDTAAYICSLATVLLACSS 187
Db 131 QPSFTSPQPPTLPPLGTGERKPFWPLHRQLPNSTVYSQRSS--HRPLCS-----SD 179

OY 188 CV 189
Db 180 CI 181

RESULT 12
MMHUB1
laminin beta-1 chain precursor - human
N:Alternate names: laminin chain B1
C:Species: Homo sapiens (man)
C>Date: 30-Jun-1991 #sequence_revision 30-Jun-1991 #text_change 21-Aug-1998
C:Accession: S13547; A28483; A26994
R:Pikkarainen, T.; Eddy, R.; Fukushima, Y.; Byers, M.; Shows, T.; Philanjaniem, T.; S.
J. Biol. Chem. 265, 10454-10462, 1987
A>Title: Human laminin B1 chain. A multidomain protein with gene (LAMBI) locus in the
A:Reference number: A28483; MUID:87280097
A:Accession: A28483
A:Molecule type: DNA
A:Residues: 1-1786 <YU>
A:Cross-references: GB:M61951; GB:J02778; NID:g186911; PID:g186913
A>Note: The nucleotide sequence was submitted to Genbank, February 1991
R:Pikkarainen, T.; Eddy, R.; Fukushima, Y.; Byers, M.; Shows, T.; Philanjaniem, T.; S.
J. Biol. Chem. 265, 10454-10462, 1987
A>Title: Human laminin B1 chain. A multidomain protein with gene (LAMBI) locus in the
A:Reference number: A28483; MUID:87280097
A:Accession: A28483
A:Molecule type: mRNA
A:Residues: 1-1786 <PIK>
A:Cross-references: GB:M61951; GB:J02778; NID:g186911; PID:g186913
R:Jaye, M.; Modi, W.S.; Ricca, G.A.; Mudd, R.; Chiu, I.M.; O'Brien, S.J.; Drohan, W.N.N
. Hum. Genet. 41, 605-615, 1987
A>Title: Isolation of a cDNA clone for the human laminin-B1 chain and its gene localiza
A:Reference number: A26994; MUID:88021029
A:Accession: A26994
A:Molecule type: mRNA
A:Residues: 1276-1469, 'V', 1471-1695, 'G', 1697-1709 <JAN>
A:Cross-references: EMBL:M20206; NID:g186914; PID:g186915
C:Genetics:
A:Gene: GDB:LAMB1
A:Cross-references: GDB:119357; OMIM:150240
A:Map position: 7q31.1-7q31.3
A:Intons: 13/1; 71/3; 117/1; 141/3; 204/3; 226/1; 293/3; 334/1; 397/1; 457/1; 494/3;
64/3; 1513/1; 1582/2; 1629/3; 1688/3; 1742/1
C:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type lamin
C:Function:
A>Description: Interact with cells and with other basement membrane proteins to promo
C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellu
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-1786/Product: laminin beta-1 chain #status predicted <MAT>

A: Molecule type: protein
A: Residues: 457-466;854-868;932-946 <HAR>
R: Mann, K.; Deutzmann, R.; Timpl, R.
Eur. J. Biochem. 178, 71-80, 1988
A: Title: Characterization of proteolytic fragments of the laminin-nidogen complex and the
A: Reference number: S08895; MUID:89078415
A: Accession: S14877
A: Molecule type: protein
A: Residues: 590-620 <MAN>
R: Barlow, D.P.; Green, N.M.; Kurkinen, M.; Hogan, B.L.M.
EMBO J. 3, 2355-2362, 1984
A: Title: Sequencing of laminin B chain cDNAs reveals C-terminal regions of collod-coll
A: Reference number: A02870; MUID:85051302
A: Accession: A02871
A: Molecule type: mRNA
A: Residues: 1292-1530, 'MBMP', 1535-1691, 'C', 1693-1748, 'N', 1750-1786 <BAR>
A: Cross-references: EMBL:X05212; NID:G52861; PID:G809042
R: Deutzmann, R.; Huber, J.; Schmetz, K.A.; Oberbauer, I.; Hartl, L.
Eur. J. Biochem. 177, 35-45, 1988
A: Title: Structural study of long arm fragments of laminin. Evidence for repetitive C-ter
A: Reference number: S01790; MUID:89030693
A: Accession: S02036
A: Molecule type: protein
A: Residues: 1561-1587 <DEU>
R: Paulsson, M.; Deutzmann, R.; Timpl, R.; Dalzoppo, D.; Odermatt, E.; Engel, J.
EMBO J. 4, 309-316, 1985
A: Title: Evidence for collod-coll alpha-helical regions in the long arm of laminin.
A: Reference number: S13543; MUID:85257455
A: Accession: S13543
A: Molecule type: protein
A: Residues: 1700-1748, 'N', 1750-1759 <PAU>
C: Genetics:
A: Gene: Lamb-1
A: Map position: 12
C: Complex: laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin
C: Function:
A: Description: interact with cells and with other basement membrane proteins to promote
C: Superfamily: laminin beta-1 chain; laminin-type EGF-like homology
C: Keywords: basement membrane; calcium binding; cell binding; collod coll; extracellular
F: 1-21/Domain: signal sequence #status predicted <SIG>
F: 22-1786/Product: laminin beta-1 chain #status predicted <MAT>
F: 22-270/Domain: VI <DOM5>
F: 271-540/Domain: V <DOM5>
F: 271-332/Domain: laminin-type EGF-like homology <EG1>
F: 335-395/Domain: laminin-type EGF-like homology <EG2>
F: 398-455/Domain: laminin-type EGF-like homology <EG3>
F: 458-507/Domain: laminin-type EGF-like homology <EG4>
F: 510-540/Domain: laminin-type EGF-like homology #status atypical <EG5>
F: 541-712/Domain: IV <DOM4>
F: 773-1182/Domain: III <DOM3>
F: 773-818/Domain: laminin-type EGF-like homology <EG6>
F: 821-864/Domain: laminin-type EGF-like homology <EG7>
F: 867-914/Domain: laminin-type EGF-like homology <EG8>
F: 917-973/Domain: laminin-type EGF-like homology <EG9>
F: 976-1025/Domain: laminin-type EGF-like homology <EG10>
F: 1028-1081/Domain: laminin-type EGF-like homology <EG11>
F: 1084-1129/Domain: laminin-type EGF-like homology <EG12>
F: 1132-1176/Domain: laminin-type EGF-like homology <EG13>
F: 1183-1397/Domain: II, heptad repeats <DOM2>
F: 1398-1430/Domain: alpha <ALP>
F: 1431-1786/Domain: I, heptad repeats <DOM1>
F: 22/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predicted
F: 30/Disulfide bonds: #status predicted
F: 120-336, 519-677, 1041-1195, 1279, 1336, 1343, 1487, 1533, 1542, 1643/Binding site: carboxylat
F: 1179, 1182, 1785/Disulfide bonds: interchain #status predicted

Db 756 RNIIISALIHOTGLACE---CDPQGSLSVCDPNGGQCQCRRNRYGRTCNRCAPRG-- 811

QY 59 SKCEGFGYGEDAQCVCPCRPNR-----FKEDMGFGKC 89

Db 811 -----FGRGPMNG-CRKCDCCHLGQASAFCAFDATIGQCHCFQGIYARQCDCRCLPGYWGVPSC 864

QY 90 KPC-----ADCALYNRQRANMCHTSDAVCGDCLPGFYRTKVLGVFDQDMCVCCGDP 143

Db 865 QPCQCGHAGLDCDVTGECCLSCODYTTGHCNCRCLAGY-GDPIIGSGD-HCRPCPCPDG 922

QY 144 P 144

Db 923 P 923

RESULT 15

GOHUN

nerve growth factor receptor precursor, low affinity - human

N:Alternate names: NGF receptor

C:Species: Homo sapiens (man)

C:Date: 31-Mar-1998 #sequence_revision 31-Mar-1998 #text_change 01-May-1998

C:Accession: A25218; A60204; S21689; I57638

R:Johnson, D.; Lamanah, A.; Buck, C.R.; Sehgal, A.; Morgan, C.; Mercer, E.; Bothwell, Cell 47, 545-554, 1986

A:Title: Expression and structure of the human NGF receptor.

A:Reference number: A25218; MUID:87051725

A:Accession: A25218

A:Molecule type: mRNA

A:Residues: 1-427 <JOH>

A:Cross-references: GB:M14764; NID:g189204; PID:g189205

R:Marano, N.; Dietzschold, B.; Earley Jr., J.J.; Schattman, G.; Thompson, S.; Grob, J. Neurochem. 48, 225-233, 1987

A:Title: Purification and amino terminal sequencing of human melanoma nerve growth fa

A:Reference number: A60204; MUID:87085574

A:Accession: A60204

A:Molecule type: Protein

A:Residues: 29-31,'T',33-42,'TT',45-46,'TX',50-51,'XX',54-56 <MAR>

A:Experimental source: melanoma cell line A875

A>Note: This sequence has been corrected by a note added in proof to follow the nucle

R:Visavajjalala, P.; Leszyk, J.D.; Iln-Goeke, J.; Ross, A.H. Arch. Biochem. Biophys. 294, 244-252, 1992

A:Title: Structural domains of the extracellular domain of human nerve growth factor

A:Reference number: S21689; MUID:92198017

A:Accession: S21689

A>Status: preliminary

A:Molecule type: protein

A:Residues: 185-208 <YIS>

R:Sehgal, A.; Patil, N.; Chao, M. Mol. Cell. Biol. 8, 3160-3167, 1988

A:Title: A constitutive promoter directs expression of the nerve growth factor recept

A:Reference number: I57638; MUID:89096903

A:Accession: I57638

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-22 <RES>

A:Cross-references: GB:M21621; NID:g189206; PID:g189207

C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastom

C:Comment: The cysteine-rich region of the extracellular domain may form part or all

C:Comment: This protein is thought to form a high-affinity receptor when it associate

C:Comment: This receptor undergoes both N- and O-linked glycosylation.

C:Genetics:

A:Gene: GDB:NGFR

A:Cross-references: GDB:120234; OMIM:162010

A:Map position: 17q21-17q22

C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology

C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-427/Product: nerve growth factor receptor #status experimental <MNT>

F:29-250/Domain: extracellular #status predicted <XT>

F:32-65/Domain: NGF receptor repeat homology <NLI>

F:67-108/Domain: NGF receptor repeat homology <NG2>

F:109-147/Domain: NGF receptor repeat homology <NG3>

F:149-189/Domain: NGF receptor repeat homology <NG4>


```

ID OX40_MOUSE STANDARD; PRT; 272 AA.
AC P47741;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE OX40L RECEPTOR PRECURSOR (OX40 ANTIGEN).
OS TXGP1 OR OX40.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE; 94044750.
RA CALDERHEAD D.M., BUELMANN J.E., VAN DEN EERTWEGH A.J.,
RA CLAASSEN E., NOELLE R.J., FELL H.;
RL J. IMMUNOL. 151:5261-5271(1993).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE; 95255413.
RA BIRKELAND M.L., COPELAND N.G., GILBERT D.J., JENKINS N.A.,
RA BARCLAY A.N.;
RL EUR. J. IMMUNOL. 25:926-930(1995).
CC -1- FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; Z21674; G312828; -.
DR MGI; X85214; G732819; -.
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS50050; TNFR_NGFR_2; 2.
KW RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT;
KW SIGNAL.
FT CHAIN 1 19 POTENTIAL.
FT DOMAIN 20 272 OX40L RECEPTOR.
FT TRANSMEM 20 211 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 212 236 POTENTIAL.
FT DOMAIN 237 272 CYTOPLASMIC (POTENTIAL).
FT REPEAT 26 165 4 X TNFR-CYS.
FT REPEAT 61 103 TNFR-CYS 1.
FT REPEAT 104 124 TNFR-CYS 2.
FT REPEAT 125 165 TNFR-CYS 4.
FT CARBOHYD 144 144 POTENTIAL.
FT CONFLICT 15 15 A -> G (IN REF. 2).
SQ SEQUENCE 272 AA; 30153 MW; 67D1B978 CRC32;

```

Query Match 13.3%; Score 155; DB 1; Length 272;
Best Local Similarity 27.6%; Pred. No. 1.4e-06;
Matches 59; Conservative 27; Mismatches 88; Indels 40; Gaps 14;

```

OY 18 LFLHLACKVSCFEGDCROQEFKRSNG-CYLCROCGFGMELSECGFGYEDACQVPCR 76
DB 11 LLLALTLGLVAPRLNCKVHTYP--SGHKC--CRECOPGHGMSVRC--DHTRDILCHPCE 64
OY 77 PHREKEDMGFOCKPCADCAVNFQ--RANGSHSDAVCGSLGFGYKTKLVGFQD--- 133
DB 65 TGFNEAVNVTCTCKOCTOCNHRSGSELKONCTPQDVTVC-RCRPQTOPR-----QDSGY 117
OY 133 ---MECVPCGDPPEYEP---HCTSKVNLKISSTVSSPDTFAVAVIC---SALATVL 182
DB 118 KLGVDCVPC--PQGHFSPGNGNACKPMTNCLSGKQTHRPSDSLDAV-CEDRSLATL 174
OY 183 LACS-----SCVSTARGSSWRNPVSSHPSV 210
DB 175 WETQRPTEPFTVOST---TVMPTSTSLSPBPTL 205

```

RESULT 3
TNRC_HUMAN
ID TNRC_HUMAN STANDARD; PRT; 435 AA.
AC P36941;

```

DT 01-JUN-1994 (REL. 29, CREATED)
DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
DE 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR)
DE 2 RELATED PROTEIN (TUMOR NECROSIS FACTOR C RECEPTOR).
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE; 93252381.
RA BAENS M., CHAFFANER M., CASSIMAN J.J., DEN BERGHE H., MARVEN P.;
RL GENOMICS 16:214-218(1993).
RN [2]
RP FUNCTION: 94225209.
RA CROME P.D., VAN ARSDALE T.L., WALTER B.N., WARE C.F., HESSON C.,
RA EHRENFELS B., BROWNING J.L., DIN W.S., GOODWIN R.G., SMITH C.A.;
RL SCIENCE 264:707-710(1994).
CC -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
CC -1- IMMUNE DEVELOPMENT.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; I04270; G339762; -.
DR HSP; P19999; ICLG.
DR MIM; 600979; -.
DR PROSITE; PS00652; TNFR_NGFR_1; 2.
DR PROSITE; PS50050; TNFR_NGFR_2; 3.
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
KW SIGNAL.
FT CHAIN 1 30 POTENTIAL.
FT DOMAIN 31 435 LYMPHOTOXIN-BETA RECEPTOR.
FT TRANSMEM 31 227 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 228 248 POTENTIAL.
FT DOMAIN 249 435 CYTOPLASMIC (POTENTIAL).
FT REPEAT 42 211 4 X TNFR-CYS.
FT REPEAT 82 124 TNFR-CYS 1.
FT REPEAT 125 168 TNFR-CYS 2.
FT REPEAT 169 211 TNFR-CYS 3.
FT REPEAT 199 211 TNFR-CYS 4.
FT DISULFID 43 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 80 BY SIMILARITY.
FT DISULFID 83 98 BY SIMILARITY.
FT DISULFID 101 116 BY SIMILARITY.
FT DISULFID 104 124 BY SIMILARITY.
FT DISULFID 126 132 BY SIMILARITY.
FT DISULFID 139 148 BY SIMILARITY.
FT DISULFID 142 167 BY SIMILARITY.
FT DISULFID 170 185 BY SIMILARITY.
FT CARBOHYD 40 40 POTENTIAL.
FT CARBOHYD 177 177 POTENTIAL.
SQ SEQUENCE 435 AA; 46709 MW; 203B82DD CRC32;

```

Query Match 12.1%; Score 140.5; DB 1; Length 435;
Best Local Similarity 24.1%; Pred. No. 3.4e-05;
Matches 57; Conservative 29; Mismatches 74; Indels 77; Gaps 16;

```

OY 30 ETGCGROE---FKDRSGNCVLCROGPGMELSECGFGYEDACQVPCRPHREKEDMG 86
DB 39 ENQTCRQDEKEYEYEPQHRIC--CSRCPGYYSKAC--SRIRDTVCATCAENSTNEHWNY 94
OY 87 ---QKCRPC-----ADCAVNFQ-----RANGSHSDAVCGDCLPGF 121
DB 95 LTIQQLRPPDVPVGLFEIAPCTSKRTQCRQPGMCAMALECTHCE--LLSDCPPT 152
OY 122 YRTRK-LVGFQDMECVPC-----GDPPEYEPH--CTSKVNLKIS-----STVS 163
DB 153 EAELEKDEVGKGNHCVCAGKHGFPONTSSPSARCOPHTRCENO-GLVABAGTAQSDTTCK 211
OY 164 SPRD-----TAVAVICGALATVILACS---SCVSTARGSSWRNPVSSHPSV 210

```



```

FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 454 TUMOR NECROSIS FACTOR RECEPTOR 1.
FT DOMAIN 22 212 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 213 235 POTENTIAL.
FT DOMAIN 236 454 CYTOSOLASMIC (POTENTIAL).
FT DOMAIN 43 196 4 X TNFR-CYS.
FT REPEAT 43 82 TNFR-CYS 1.
FT REPEAT 83 125 TNFR-CYS 2.
FT REPEAT 126 166 TNFR-CYS 3.
FT REPEAT 167 196 TNFR-CYS 4.
FT DOMAIN 339 349 N-SMASE ACTIVATION DOMAIN (NSD).
FT DOMAIN 356 441 DEATH DOMAIN.
FT DISULFID 44 58 BY SIMILARITY.
FT DISULFID 59 72 BY SIMILARITY.
FT DISULFID 62 81 BY SIMILARITY.
FT DISULFID 84 99 BY SIMILARITY.
FT DISULFID 102 117 BY SIMILARITY.
FT DISULFID 105 125 BY SIMILARITY.
FT DISULFID 127 143 BY SIMILARITY.
FT DISULFID 146 158 BY SIMILARITY.
FT DISULFID 149 166 BY SIMILARITY.
FT DISULFID 168 179 BY SIMILARITY.
FT DISULFID 182 191 BY SIMILARITY.
FT DISULFID 185 195 BY SIMILARITY.
FT CARBOHYD 54 54 POTENTIAL.
FT CARBOHYD 151 151 POTENTIAL.
FT CARBOHYD 202 202 POTENTIAL.
FT CONFLICT 394 394 R -> G (IN REF. 6).
SQ SEQUENCE 454 AA; 50129 MM; 4B6BEC09 CRC32;

```

```

Query Match 11.9%; Score 138.5; DB 1; Length 454;
Best Local Similarity 22.0%; Pred. No. 5.3e-05;
Matches 56; Conservative 31; Mismatches 103; Indels 65; Gaps 11;

```

```

QY 6 LPLHRTVLFALFLFLHLLACKVSCETG-----DCRQOEKDRSGNCVLCQK 52
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 3 LPTVPGLLSLVLALLMGHPGVTGLVPSLDREKRDSLCPQKGVHKKNSICCTKC 62
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 53 GPGMELSKEGFGYGDACVPCPRHREKEDWGF-QKCKFCADCAL-VNFRQANCSHTS 110
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 63 HKGTLYVSDCP-SPGRDTCRECEKGTFTASQVYLROCLSKTCRKMSQVEISPCQADK 121
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 111 DAVCG-----DCLDGEFRKTKL-----VGF-QDMECVP 137
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 122 DTYGCGCKENOFRLSTHQCVCSCFENGVTITPKETONTVCNCHAGFPLRESECV 181
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 138 CGDPPPEPHCTSKVLVVIS-----STVSSPRDTVAVAVICSALATVLLACSSC-VSS 191
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 182 CS-----HCKKNBECMCLCPPLANTYNTNQDSTAVLLPLVILLGLCLLSFIFISL 233
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 192 TANGSSWRNPVAVSS 206
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 234 MCRYPRWR--PEVYS 246
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
VC22_VARY STANDARD; PRT; 349 AA.
AC P34015;
DT 01-FEB-1994 (REL. 28, CREATED)
DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
DE 01-JUN-1994 (REL. 29, LAST ANNOTATION UPDATE)
DE PROTEIN C22/B28 HOMOLOG.
GN G4R.
OS VARIDOLA VIRUS.
OC VIRIDAE: DS-DNA ENVELOPED VIRUSES; POXYVIRIDAE: CHOROPHOXVIRINAE;
OC ORTHOPOXYVIRUSES.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-INDIA-1967 / ISOLATE IND3;
RX MEDLINE: 93202281.
RA SHCHELKUNOV S.N., BLINOV V.M., SANDAKHCHIEV L.S.;

```

```

RL FEBS LETT. 319:80-83(1993).
CC -1 SIMILARITY: CONTAINS 2 LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGIONS.
DR EMBL: X69198; G457087; -.
DR EMBL: X67117; G516449; -.
DR PIR: D36858; D36858.
DR PIR: S35987; S35987.
DR PIR: S46888; S46888.
DR HSSP: p19438; 1TNFR.
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
DR PROSITE: PS50050; TNFR_NGFR_2; 2.
KW REPEAT.
FT DOMAIN 31 108 2 X TNFR-CYS.
FT REPEAT 31 66 TNFR-CYS 1.
FT REPEAT 67 108 TNFR-CYS 2.
SQ SEQUENCE 349 AA; 38189 MM; 50D0B435 CRC32;

```

```

Query Match 11.9%; Score 138.5; DB 1; Length 349;
Best Local Similarity 23.7%; Pred. No. 4.2e-05;
Matches 50; Conservative 30; Mismatches 88; Indels 43; Gaps 10;

```

```

QY 10 RTVLFALFLFLHLL-----ACKVSCETGDCRQOEKDRSGNCVLCQKQGPMELSKECG 63
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 2 KSVLYLYILFSLCIIINGRDAAPYTPPNCKCKDTEYK-RHNLG-CISCPETVSRIC- 58
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 64 FGYGDAQCVPCPRHRE-KEDMGFOKCKPCADCALVNFRQANCSHTSDAVCGDCLPGFY 122
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 58 -DSKNTVQCTPCGSGTFTSRNNHLPALCLISNGRCSNNOVETRSCTTHNRIC-ECSPGY 115
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 123 -----RKTKL-VGF-----QDMECVPG-----DPPPEPHCTSK 152
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 116 CLKKSGSGKACVSOIKCGIGYVSGHTSVGVICSPCGFTYSHRTVSSADKCEVPNNT 175
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 153 VMLVKISSTVSSPRDTVAVAVICSALATVLL 183
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 176 FNVIDEITLYPVNDISCTRTTTGLSESL 206
   || : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 7
NGFR_CHICK STANDARD; PRT; 416 AA.
ID NGFR_CHICK
AC P18519;
DT 01-NOV-1990 (REL. 16, CREATED)
DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)
DE (GP80-LINGFR) (P75 ICD).
GN NGFR.
OS GALLUS GALLUS (CHICKEN)
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; AVES; NEOGNATHAE;
OC GALLIFORMES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN.
RX MEDLINE: 90166579.
RA LARGE T.H., WESKAMP G., HEIDER J.C., RADEKE M.J., MISKO T.P.,
RA SHOOTER E.M., REICHARDT L.F.;
RL NEURON 2:1123-1134(1989).
[2]
RP SEQUENCE OF 21-416 FROM N.A.
RX MEDLINE: 90152140.
RA HEUER J.G., FATEMIE-NAINE S., WHEELER E.F., BOTHWELL M.;
RL DEV. BIOL. 137:287-304(1990).
CC -1 FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
CC NT-3, AND NT-4.
CC -1 SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
CC BOND FORMATION.
CC -1 SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1 PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
CC -1 SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR PIR: JN0006; JN0006.
DR PIR: A60504; A60504.
DR PROSITE: PS00652; TNFR_NGFR_1; 3.

```



```
FT DISULFID 288 313 BY SIMILARITY.
FT DISULFID 315 324 BY SIMILARITY.
FT DISULFID 327 347 BY SIMILARITY.
FT DISULFID 350 359 BY SIMILARITY.
FT DISULFID 352 377 BY SIMILARITY.
FT DISULFID 380 389 BY SIMILARITY.
FT DISULFID 392 410 BY SIMILARITY.
FT DISULFID 413 426 BY SIMILARITY.
FT DISULFID 415 441 BY SIMILARITY.
FT DISULFID 443 452 BY SIMILARITY.
FT DISULFID 455 470 BY SIMILARITY.
FT DISULFID 473 487 BY SIMILARITY.
FT DISULFID 475 494 BY SIMILARITY.
FT DISULFID 496 505 BY SIMILARITY.
FT DISULFID 508 522 BY SIMILARITY.
FT DISULFID 784 796 BY SIMILARITY.
FT DISULFID 786 803 BY SIMILARITY.
FT DISULFID 805 814 BY SIMILARITY.
FT DISULFID 817 829 BY SIMILARITY.
FT DISULFID 832 844 BY SIMILARITY.
FT DISULFID 834 851 BY SIMILARITY.
FT DISULFID 853 862 BY SIMILARITY.
FT DISULFID 865 875 BY SIMILARITY.
FT DISULFID 878 887 BY SIMILARITY.
FT DISULFID 880 894 BY SIMILARITY.
FT DISULFID 897 906 BY SIMILARITY.
FT DISULFID 909 925 BY SIMILARITY.
FT DISULFID 928 944 BY SIMILARITY.
FT DISULFID 930 955 BY SIMILARITY.
FT DISULFID 957 966 BY SIMILARITY.
FT DISULFID 969 984 BY SIMILARITY.
FT DISULFID 987 1001 BY SIMILARITY.
FT DISULFID 989 1008 BY SIMILARITY.
FT DISULFID 1011 1020 BY SIMILARITY.
FT DISULFID 1023 1036 BY SIMILARITY.
FT DISULFID 1096 1108 BY SIMILARITY.
FT DISULFID 1098 1115 BY SIMILARITY.
FT DISULFID 1117 1126 BY SIMILARITY.
FT DISULFID 1129 1141 BY SIMILARITY.
FT DISULFID 1144 1156 BY SIMILARITY.
FT DISULFID 1146 1163 BY SIMILARITY.
FT DISULFID 1165 1174 BY SIMILARITY.
FT DISULFID 1177 1188 BY SIMILARITY.
FT DISULFID 1191 1191 INTERCHAIN (PROBABLE).
FT DISULFID 1194 1194 INTERCHAIN (PROBABLE).
FT DISULFID 1798 1798 INTERCHAIN (PROBABLE).
FT CARBOHYD 251 251 POTENTIAL.
FT CARBOHYD 371 371 POTENTIAL.
FT CARBOHYD 1086 1086 POTENTIAL.
FT CARBOHYD 1250 1250 POTENTIAL.
FT CARBOHYD 1309 1309 POTENTIAL.
FT CARBOHYD 1349 1349 POTENTIAL.
FT CARBOHYD 1500 1500 POTENTIAL.
SQ SEQUENCE 1799 AA; 196352 MW; AF310037 CRC32;
```

```
Query Match 11.2%; Score 130.5; DB 1; Length 1799;
Best Local Similarity 23.9%; Pred. No. 0.00081;
Matches 55; Conservative 22; Mismatches 84; Indels 69; Gaps 14;
```

```
OY 7 PLHRT-----VLAAILFLHLACKVCETGDCRQOEFKDRSGNCVLCCKGCGKELSK 61
DB 761 PLSEFCALLISVSLIYNGALPCQ--CDPGSLSECSPHHGQC---KCKPGV-VGRR 813
OY 62 C-----GFG-----YGEDACVPCPRPHF-----KEDMGF 86
DB 814 CDVCAATGYGFGPACQACQCSPDGALSALCEGTSGQC-PCRPAGFGLRCDHCGQSGMGF 872
OY 87 QKCRPC-----ADCALVNRFOQANC-SHRSDAVCGCDLPGFTRKTKL-VGFODMECVPCG 139
DB 873 PNCRCVCNGRDECDTHGACLGCRDYGGEHCRCIAGFGDPLPYGGGCRPC-PCP 931
OY 140 DPPPPEPCHTSKVNLIKISSVSSPRDTAVAAVICSALATVL-LACSSC 188
```

```
Db 932 EGPGSQRRHFAI-----SCHRDGYSQQLVCHCRACVYGLRCEAC 969
RESULT 9
TNRL_RAT TNRL_RAT STANDARD; PRT; 461 AA.
AC P22934;
DT 01-AUG-1991 (REL. 19, CREATED)
DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR 1 PRECURSOR (P60) (TNF-R1) (P55).
GN TNFR1 OR TNFR-1.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUKARYOTA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 91090841.
RA HIMMLER A., MAUER-FOGY I., KROENKE M., SCHEURICH P., PRIZENMAIER K.,
RA LANTZ M., OLSSON I., HAUPTMANN R., STRATOWA C., ADOLF G.R.;
RL DNA CELL BIOL. 9:705-715(1990).
CC -I- FUNCTION: RECEPTOR FOR TNF-ALPHA. THE ADAPTOR MOLECULE FADD
CC RECRUITS CASPASE-8 TO THE ACTIVATED RECEPTOR. THE RESULTING
CC AGGREGATE CALLED THE DEATH-INDUCING SIGNALING COMPLEX (DISC)
CC PERFORMS CASPASE-8 PROTEOLYTIC ACTIVATION WHICH INITIATES THE
CC SUBSEQUENT CASCADE OF CASPASES (ASPARTATE-SPECIFIC CYSTEINE
CC PROTEASES) MEDIATING APOPTOSIS (BY SIMILARITY).
CC -I- SUBUNIT: TNF BINDING TO THE EXTRACELLULAR DOMAIN OF TNFR1 LEADS TO
CC HOMOTRIMERIZATION. ONCE AGGREGATED THE RECEPTORS DEATH DOMAINS
CC PROVIDE A NOVEL MOLECULAR INTERFACE THAT INTERACTS SPECIFICALLY
CC WITH THE DEATH DOMAIN OF TRADD. VARIOUS TRADD-INTERACTING
CC PROTEINS SUCH AS TRAFS, RIP AND POSSIBLY FADD, ARE RECRUITED TO
CC TNFR1 COMPLEX BY THEIR ASSOCIATION WITH TRADD. THIS COMPLEX
CC ACTIVATES AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
CC NF-KAPPA B SIGNALING (BY SIMILARITY).
CC -I- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
CC EMBL; M63122; G207362; -.
DR PIR; B36355; B36355.
DR HSSP; P19438; TNFR.
DR PROSITE; PS00652; TNFR_NGFR.1; 3.
DR PROSITE; PS50050; TNFR_NGFR.2; 3.
DR PROSITE; PS50017; DEATH_DOMAIN: 1.
KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL; APOPTOSIS.
FT SIGNAL 1 21
FT CHAIN 1 21
FT DOMAIN 22 455
FT TRANSMEM 212 234
FT DOMAIN 235 461
FT DOMAIN 43 196
FT REPEAT 43 82
FT REPEAT 83 125
FT REPEAT 126 166
FT REPEAT 167 196
FT DOMAIN 344 354
FT DOMAIN 363 448
FT DISULFID 44 58
FT DISULFID 59 72
FT DISULFID 62 81
FT DISULFID 84 99
FT DISULFID 102 117
FT DISULFID 105 125
FT DISULFID 127 143
FT DISULFID 146 158
FT DISULFID 149 166
FT DISULFID 168 179
FT DISULFID 182 191
FT DISULFID 185 195
FT CARBOHYD 54 54
FT CARBOHYD 151 151
FT CARBOHYD 201 201
SQ SEQUENCE 461 AA; 50969 MW; 82F6B08 CRC32;
```


FT TRANSMEM 183 203 POTENTIAL.
 FT DOMAIN 204 250 CYTOPLASMIC (POTENTIAL).
 FT REPEAT 26 141 3 X TNR-CYS.
 FT REPEAT 64 104 TNR-CYS 1.
 FT REPEAT 105 141 TNR-CYS 2.
 FT CARBOHYD 95 95 TNR-CYS 3.
 FT CARBOHYD 162 162 POTENTIAL.
 SQ SEQUENCE 250 AA; 28164 MW; AF6F74ED CRC32.

Query Match 10.88; Score 125.5; DB 1; Length 250;
 Best Local Similarity 26.9%; Pred. No. 0.00041;
 Matches 49; Conservative 21; Mismatches 59; Indels 53; Gaps 12;

OY 34 CROGEFKDRSGNCVLCQCGPQGMELSKCGRGYED---AQCVCRP-HRFEKMGFQ-K 88
 DB 27 CPDHWHTGGGLC--CMWCEGTFEVRDC---EODPTAACDCPCIGTSFSPDYHTRPH 80
 OY 89 CKPCADCALVNRQFANCSTSDAVCGDCLGFYRKTKLGVGFODMECVPCG----- 140
 DB 81 CESGRHCN--SGFLIRNCIYTANEC--SC-----SKNQCROEECTEDCDPLNPLTR 130
 OY 140 -----DPPPEYEPCTSK---VNLVKISSTVSSPRDTAAVAATCSALATVLLACSS 187
 DB 131 QPSETSPQPPPHLPHTGTEKPSWPLHRLQPNSTVYSQFSS--HRPLCS-----SD 179
 OY 188 CV 189
 DB 180 CT 181

RESULT 12

LMB1.HUMAN STANDARD; PRT; 1786 AA.

ID LMB1.HUMAN
 AC P07942;
 DT 01-AUG-1988 (REL. 08, CREATED)
 DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN).
 GN LAMB1.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RX SEQUENCE FROM N.A.
 RA VOULTEENAHO R.; CHOW L.T.; TRYGGVASON K.;
 RL J. BIOL. CHEM. 265:15611-15616(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 87280097.
 RA PIKKARAINEN T.; EDDY R.; FUKUSHIMA Y.; BYERS M.; SHOWS T.,
 RA PIHLAJANTIE T.; SARASTE M.; TRYGGVASON K.;
 RL J. BIOL. CHEM. 262:10454-10462(1987).
 RN [3]
 RP SEQUENCE OF 1276-1709 FROM N.A.
 RX MEDLINE: 88021029.
 RA JAYE M.; MODI W.S.; RICCA G.A.; MUDD R.; CHIU I.M.; O'BRIEN S.J.,
 RA BROHAN W.N.;
 RL AM. J. HUM. GENET. 41:605-615(1987).
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
 IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
 CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
 WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
 DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
 TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
 COMPRISING ONE LONG & THREE SHORT ARMS WITH CLOUTELS AT EACH END.
 CC -1- SUBUNIT: THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ),
 LAMININ-2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR

CC COMPONENT).
 CC -1- WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
 CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
 CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
 CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
 DR EMBL: M51951; G186913; JOINED.
 DR EMBL: M51917; G186913; JOINED.
 DR EMBL: M61918; G186913; JOINED.
 DR EMBL: M61921; G186913; JOINED.
 DR EMBL: M61922; G186913; JOINED.
 DR EMBL: M61923; G186913; JOINED.
 DR EMBL: M61924; G186913; JOINED.
 DR EMBL: M61925; G186913; JOINED.
 DR EMBL: M61926; G186913; JOINED.
 DR EMBL: M61927; G186913; JOINED.
 DR EMBL: M61928; G186913; JOINED.
 DR EMBL: M61929; G186913; JOINED.
 DR EMBL: M61930; G186913; JOINED.
 DR EMBL: M61931; G186913; JOINED.
 DR EMBL: M61932; G186913; JOINED.
 DR EMBL: M61933; G186913; JOINED.
 DR EMBL: M61934; G186913; JOINED.
 DR EMBL: M61935; G186913; JOINED.
 DR EMBL: M61936; G186913; JOINED.
 DR EMBL: M61938; G186913; JOINED.
 DR EMBL: M61939; G186913; JOINED.
 DR EMBL: M61940; G186913; JOINED.
 DR EMBL: M61941; G186913; JOINED.
 DR EMBL: M61942; G186913; JOINED.
 DR EMBL: M61943; G186913; JOINED.
 DR EMBL: M61944; G186913; JOINED.
 DR EMBL: M61945; G186913; JOINED.
 DR EMBL: M61946; G186913; JOINED.
 DR EMBL: M61947; G186913; JOINED.
 DR EMBL: M61948; G186913; JOINED.
 DR EMBL: M61949; G186913; JOINED.
 DR EMBL: M61950; G186913; JOINED.
 DR EMBL: M55370; G186876; JOINED.
 DR EMBL: M55378; G186876; JOINED.
 DR EMBL: M55365; G186876; JOINED.
 DR EMBL: M55371; G186876; JOINED.
 DR EMBL: M55372; G186876; JOINED.
 DR EMBL: M55373; G186876; JOINED.
 DR EMBL: M55374; G186876; JOINED.
 DR EMBL: M55375; G186876; JOINED.
 DR EMBL: M55376; G186876; JOINED.
 DR EMBL: M55344; G186876; JOINED.
 DR EMBL: M55345; G186876; JOINED.
 DR EMBL: M55346; G186876; JOINED.
 DR EMBL: M55347; G186876; JOINED.
 DR EMBL: M55348; G186876; JOINED.
 DR EMBL: M55349; G186876; JOINED.
 DR EMBL: M55350; G186876; JOINED.
 DR EMBL: M55351; G186876; JOINED.
 DR EMBL: M55352; G186876; JOINED.
 DR EMBL: M55353; G186876; JOINED.
 DR EMBL: M55355; G186876; JOINED.
 DR EMBL: M55356; G186876; JOINED.
 DR EMBL: M55357; G186876; JOINED.
 DR EMBL: M55358; G186876; JOINED.
 DR EMBL: M55359; G186876; JOINED.
 DR EMBL: M55360; G186876; JOINED.
 DR EMBL: M55361; G186876; JOINED.
 DR EMBL: M55362; G186876; JOINED.
 DR EMBL: M55363; G186876; JOINED.
 DR EMBL: M55364; G186876; JOINED.
 DR EMBL: M55366; G186876; JOINED.
 DR EMBL: M55367; G186876; JOINED.
 DR EMBL: M55368; G186876; JOINED.
 DR EMBL: M55369; G186876; JOINED.

DR EMBL; M61916; G186837; -;
 DR EMBL; M20206; G186915; -;
 DR PIR; S13547; MMHUB1.
 DR MIM; 150240; -;
 DR PROSITE; PS00022; EGF_1; 9.
 DR PROSITE; PS01186; EGF_2; 2.
 DR PROSITE; PS01248; LAMININ_TYPE_EGF; 11.
 DR GLYCOPROTEIN; BASEMENT MEMBRANE; EXTRACELLULAR MATRIX; COILED COIL;
 KW LAMININ EGF-LIKE DOMAIN; CELL ADHESION; REPEAT; SIGNAL.
 FT SIGNAL 1 21
 FT CHAIN 22 1786 LAMININ BETA-1 CHAIN.
 FT DOMAIN 22 270 LAMININ N-TERMINAL (DOMAIN VI).
 FT DOMAIN 271 540 4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN V).
 FT DOMAIN 271 334 LAMININ EGF-LIKE 1.
 FT DOMAIN 335 397 LAMININ EGF-LIKE 2.
 FT DOMAIN 398 457 LAMININ EGF-LIKE 3.
 FT DOMAIN 458 509 LAMININ EGF-LIKE 4.
 FT DOMAIN 510 540 LAMININ EGF-LIKE 5 (INCOMPLETE).
 FT DOMAIN 541 771 LAMININ DOMAIN IV.
 FT DOMAIN 772 1178 8 X LAMININ EGF-LIKE REPEATS (DOMAIN III).
 FT DOMAIN 773 820 LAMININ EGF-LIKE 6.
 FT DOMAIN 821 866 LAMININ EGF-LIKE 7.
 FT DOMAIN 867 916 LAMININ EGF-LIKE 8.
 FT DOMAIN 917 975 LAMININ EGF-LIKE 9.
 FT DOMAIN 976 1027 LAMININ EGF-LIKE 10.
 FT DOMAIN 1028 1083 LAMININ EGF-LIKE 11.
 FT DOMAIN 1084 1131 LAMININ EGF-LIKE 12.
 FT DOMAIN 1132 1178 LAMININ EGF-LIKE 13.
 FT DOMAIN 1179 1397 DOMAIN II.
 FT DOMAIN 1398 1430 DOMAIN ALPHA.
 FT DOMAIN 1431 1786 DOMAIN I.
 FT DOMAIN 1216 1315 COILED COIL (POTENTIAL).
 FT DOMAIN 1353 1388 COILED COIL (POTENTIAL).
 FT DOMAIN 1442 1781 COILED COIL (POTENTIAL).
 FT DISULFID 271 280 BY SIMILARITY.
 FT DISULFID 273 298 BY SIMILARITY.
 FT DISULFID 300 309 BY SIMILARITY.
 FT DISULFID 312 332 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 337 362 BY SIMILARITY.
 FT DISULFID 365 374 BY SIMILARITY.
 FT DISULFID 377 395 BY SIMILARITY.
 FT DISULFID 398 411 BY SIMILARITY.
 FT DISULFID 400 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 440 455 BY SIMILARITY.
 FT DISULFID 458 472 BY SIMILARITY.
 FT DISULFID 460 479 BY SIMILARITY.
 FT DISULFID 481 490 BY SIMILARITY.
 FT DISULFID 493 507 BY SIMILARITY.
 FT DISULFID 773 785 BY SIMILARITY.
 FT DISULFID 775 792 BY SIMILARITY.
 FT DISULFID 794 803 BY SIMILARITY.
 FT DISULFID 806 818 BY SIMILARITY.
 FT DISULFID 821 833 BY SIMILARITY.
 FT DISULFID 823 840 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 864 BY SIMILARITY.
 FT DISULFID 867 876 BY SIMILARITY.
 FT DISULFID 869 883 BY SIMILARITY.
 FT DISULFID 886 895 BY SIMILARITY.
 FT DISULFID 914 914 BY SIMILARITY.
 FT DISULFID 917 933 BY SIMILARITY.
 FT DISULFID 919 944 BY SIMILARITY.
 FT DISULFID 946 955 BY SIMILARITY.
 FT DISULFID 958 973 BY SIMILARITY.
 FT DISULFID 976 990 BY SIMILARITY.
 FT DISULFID 978 997 BY SIMILARITY.
 FT DISULFID 1000 1009 BY SIMILARITY.
 FT DISULFID 1012 1025 BY SIMILARITY.
 FT DISULFID 1084 1096 BY SIMILARITY.

FT DISULFID 1086 1103 BY SIMILARITY.
 FT DISULFID 1105 1114 BY SIMILARITY.
 FT DISULFID 1117 1129 BY SIMILARITY.
 FT DISULFID 1132 1144 BY SIMILARITY.
 FT DISULFID 1134 1151 BY SIMILARITY.
 FT DISULFID 1153 1162 BY SIMILARITY.
 FT DISULFID 1165 1176 BY SIMILARITY.
 FT DISULFID 1179 1179 INTERCHAIN (PROBABLE).
 FT DISULFID 1182 1182 INTERCHAIN (PROBABLE).
 FT DISULFID 1785 1785 INTERCHAIN (PROBABLE).
 FT CARBOHYD 120 120 POTENTIAL.
 FT CARBOHYD 356 356 POTENTIAL.
 FT CARBOHYD 519 519 POTENTIAL.
 FT CARBOHYD 677 677 POTENTIAL.
 FT CARBOHYD 1041 1041 POTENTIAL.
 FT CARBOHYD 1195 1195 POTENTIAL.
 FT CARBOHYD 1279 1279 POTENTIAL.
 FT CARBOHYD 1336 1336 POTENTIAL.
 FT CARBOHYD 1343 1343 POTENTIAL.
 FT CARBOHYD 1487 1487 POTENTIAL.
 FT CARBOHYD 1542 1542 POTENTIAL.
 FT CONFLICT 1470 1470 L -> V (IN REF. 3).
 FT CONFLICT 1696 1696 E -> G (IN REF. 3).
 SQ SEQUENCE 1786 AA; 198065 MW; 1D0DB223 CRC32;
 Query Match 10.8%; Score 125.5; DB 1; Length 1786;
 Best Local Similarity 25.8%; Pred. No. 0.0022;
 Matches 47; Conservative 13; Mismatches 61; Indels 61; Gaps 12;
 QY 10 RTVLEALILFLH---LACKVSCETGDCRQDEFKRSGNCV-----LCQCGSPGHEL 58
 Db 756 RNIRF-SISALHQGLACE--CDPQGSLSVCDPENGQCCQCPNVGRTCMRCAPGT-- 811
 QY 59 SKECFGEYGEDACVCPCHPR-----FKEGDFQKC 89
 Db 811 ----FGGCPSS-CRPECCHLDGVSNAFCNPYGCCHCFQGYARQCRCRLPGHMGFSC 864
 QY 90 KPCA-----DCALVNRFORANC-SHTSDAVGDCCLPGFYRTKLVLGFQDMCVCVCPDP 142
 Db 865 QPCQNGHADDDPVTG-ECINLCQDYTMGNHCRCRLAGY-GDPIIGSD-HCRCPCPD 921
 QY 143 PP 144
 Db 922 GP 923
 RESULT 13
 TNR_MOUSE STANDARD; PRT; 415 AA.
 AC P50284;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR.
 GN LTR OR TNFR.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV3; TISSUE=LUNG;
 RX MEDLINE; 96072804.
 RA FORCE W.R., WALTER B.N., HESSION C., TIZARD R., KOZAK C.A.,
 RA BROWNING J.L., WARE C.F.;
 RL J. IMMUNOL. 155:5280-5288(1995).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 96163885.
 RA NAKAMURA T., TASHIRO K., NAZAREA M., NAKANO T., SASAYAMA S.,
 RA HONDO T.;
 RL GENOMICS 30:312-319(1995).
 CC -!- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN

CC IMMUNE DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: U29173; G881621; -
 DR EMBL: U30798; G600223; -
 DR EMBL: U30798; G1061327; -
 DR MGD; MGI:104875; LTR.
 DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS50050; TNFR_NGFR_2; 3.
 KM RECEPTOR: TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
 FT SIGNAL 1 30
 FT CHAIN 31 415
 FT DOMAIN 31 223
 FT TRANSMEM 224 244
 FT DOMAIN 245 415
 FT DOMAIN 42 213
 FT REPEAT 42 81
 FT REPEAT 82 124
 FT REPEAT 125 170
 FT REPEAT 171 213
 FT DISULFID 43 58
 FT DISULFID 59 72
 FT DISULFID 62 80
 FT DISULFID 83 98
 FT DISULFID 101 116
 FT DISULFID 104 124
 FT DISULFID 126 132
 FT DISULFID 139 150
 FT DISULFID 142 169
 FT DISULFID 172 187
 FT CARBOHYD 40 40
 FT CARBOHYD 179 179
 SQ SEQUENCE 415 AA; 44956 MW; 3C5DD121 CRC32;

Query Match 10.8%; Score 125; DB 1; Length 415;
 Best Local Similarity 22.8%; Pred. No. 0.00069;
 Matches 46; Conservative 24; Mismatches 80; Indels 52; Gaps 9;

QY 47 VLCKCGPGMELEKCEGFGVGEAGACVPCPHFKEDWG-FQCKPCADALVNRQR-A 104
 Db 57 VCCSRCPGGEFVAVC--SRSDTVCKCTCHNSYNEMHNLSTCOLRCPDIVLGEFEVA 114
 QY 105 MCSHTSDAVCGDCLPGF-----YRKTLYGFO-----DMECVPC- 139
 Db 115 PCSDSKRAEC-RQPGKSCYLDNECVHCEERLVLCQPTAEAVDEIMDTVNCVPC 173
 QY 139 -----GDPPEYEPCHTSKVLVKISTVSSPRDTAAVACS-----ALATVL 182
 Db 174 PGHONTSSPRACQPH-----TRCEIQLGEVAAPGTSYSDTICKNPERGAMLLAAIL 228
 QY 183 LACSCVSTARRSSWRNPAY 204
 Db 229 SLVLFLEFTVLACAMWRHPSL 250

RESULT 14
 WSL1_HUMAN
 ID WSL1_HUMAN STANDARD; PRT; 417 AA.
 AC Q93036; Q93036; Q92983; P78515; Q99831; Q99722; P78507;
 DT 01-NOV-1997 (REL. 35, CREATED)
 DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
 DE WSL-1 PROTEIN PRECURSOR (APOPTOSIS-MEDIATING RECEPTOR DR3) (APOPTOSIS-MEDIATING RECEPTOR TRAMP) (DEATH DOMAIN RECEPTOR 3) (WSL PROTEIN) (APOPTOSIS INDUCING RECEPTOR AIR) (APO-3) (LYMPHOCYTE ASSOCIATED RECEPTOR OF DEATH) (LARD).
 DE WSL1 OR WSL OR DR3 OR DDR3.
 OS HOMO SAPIENS (HUMAN)
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN (1)
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND MUTAGENESIS.

RC TISSUE-LYMPHOID;
 RX MEDLINE: 97088617.
 RA KITSON J., RAVEN T., JIANG Y.-P., GOEDEL D.V., GILES K.M., PUN K.-T.,
 RA GRINHAM C.J., BROWN R., FARROW S.N.;
 RN NATURE 384:372-375(1996).
 RP SEQUENCE FROM N.A.
 RC TISSUE-UMBILICAL VEIN ENDOTHELIAL CELLS;
 RX MEDLINE: 97081063.
 RA CHINAIYAN A.M., O'ROURKE K., YU G.-L., LYONS R.H., GARG M.,
 RA DUAN D.R., XING L., GENTZ R., NI J., DIXIT V.M.;
 RN SCIENCE 274:990-992(1996).
 RP SEQUENCE FROM N.A.
 RA DEGLI-ESPOSTI M.A., DIN W.S., COSMAN D., SMITH C.A., GOODWIN R.G.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE-HEART;
 RA MARSTERS S.A., SHERIDAN J.P., DONAHUE C.J., PITTI R.M., GRAY C.L.,
 RA GODDARD A.D., BAUER K.D., ASHKENAZI A.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA SCREATOR G., XU X.N., OLSEN A., COMPER A., TAN R., MCMICHAEL A.,
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
 RN [6]
 RP SEQUENCE OF 4-417 FROM N.A.
 RC TISSUE-BRAIN, AND FETAL LUNG;
 RX MEDLINE: 97205335.
 RA BODMER J.-L., BURNS K., SCHNEIDER P., HOFMANN K., STEINER V.,
 RA THOME M., BORNAND T., HAHNE M., SCHROETER M., WILSON A., FRENCH L.E.,
 RL IMUNITY 6:79-88(1997).
 RN [7]
 RP SEQUENCE OF 7-417 FROM N.A.
 RC TISSUE-BRAIN;
 RA CHAUDHARY P.M., HOOD L.E.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDBJ DATA BANKS.
 CC -1- FUNCTION: INDICES APOPTOSIS AND ACTIVATES NUCLEAR FACTOR KAPPA-B (NF-KAPPA-B). DIRECTLY INTERACTS WITH TRADD ADAPTOR MOLECULE. MAY PLAY A ROLE IN REGULATING LYMPHOCYTE HOMEOSTASIS.
 CC -1- SUBUNIT: HOMODIMER. INTERACTS STRONGLY VIA THE DEATH DOMAINS WITH THE TNFR1 ASSOCIATED MOLECULE TRADD AND THE TNFR1 RECEPTOR TO ACTIVATE AT LEAST TWO DISTINCT SIGNALING CASCADES. APOPTOSIS AND NF-KAPPA B SIGNALING.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THYMOCYTES AND LYMPHOCYTES. DETECTED IN LYMPHOCYTE-RICH TISSUES SUCH AS THYMUS, COLON, INTESTINE, AND SPLEEN. ALSO FOUND IN THE PROSTATE.
 CC -1- PTM: GLYCOSYLATED (PROBABLE).
 CC -1- ALTERNATIVE PRODUCTS: THREE ISOFORMS WSL-1/LARD-1A (SHOWN HERE), WSL-S1/LARD-3, AND WSL-S2 ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: Y09382; E281462; -
 DR EMBL: Y09382; E281460; -
 DR EMBL: Y09382; E281461; -
 DR EMBL: U72763; G1669512; -
 DR EMBL: U83599; G1800297; -
 DR EMBL: U83600; G1800297; -
 DR EMBL: U78029; G1778764; -
 DR EMBL: U74611; G1763293; -
 DR EMBL: U94501; G2071949; -
 DR EMBL: U94504; G2071955; -
 DR EMBL: U75380; G1695925; -
 DR EMBL: U75381; G1695927; -
 DR EMBL: U83587; G1800293; -
 DR PROSITE: PS00652; TNFR_NGFR_1; 2.
 DR PROSITE: PS50050; TNFR_NGFR_2; 1.
 DR PROSITE: PS50017; DEATH_DOMAIN; 1.
 KW RECEPTOR; APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; SIGNAL; REPEAT.

Db 183 PMADACEEIPGRNIPRSTPEGSDSTAPSTOEPEVPEQDLV--PSTVADMVTTVMGSS 240
QY 173 -----VISCALATVLLACSSCVSTARGSSWRPNPAVSSHP 208
Db 241 QPVYTRGTNDLIPYCSILAAVVGVLVAYTAFKRNNSCKONKOGANSRP 290

Search completed: May 19, 1999, 14:43:38
Job time: 45 sec

| | | | | | | | |
|----|-------|------|-----|----|--------|--------|-----------------|
| 1 | 160.5 | 13.8 | 348 | 11 | 057277 | 057277 | monkeypox v |
| 2 | 160.5 | 13.8 | 348 | 11 | 057108 | 057108 | monkeypox v |
| 3 | 159 | 13.7 | 349 | 11 | 057291 | 057291 | monkeypox v |
| 4 | 159 | 13.7 | 349 | 11 | 057099 | 057099 | monkeypox v |
| 5 | 159 | 13.7 | 349 | 11 | 057100 | 057100 | monkeypox v |
| 6 | 159 | 13.7 | 349 | 11 | 057101 | 057101 | monkeypox v |
| 7 | 159 | 13.7 | 349 | 11 | 057102 | 057102 | monkeypox v |
| 8 | 156.5 | 13.5 | 348 | 11 | 057103 | 057103 | monkeypox v |
| 9 | 152 | 13.1 | 267 | 4 | 002764 | 002764 | oryctolagus |
| 10 | 148 | 12.7 | 350 | 11 | 057116 | 057116 | cowpox virus |
| 11 | 143.5 | 12.3 | 351 | 11 | 073559 | 073559 | cowpox virus |
| 12 | 142.5 | 12.3 | 283 | 2 | 092956 | 092956 | homo sapien |
| 13 | 142 | 12.2 | 625 | 10 | 035305 | 035305 | mus musculus |
| 14 | 141.5 | 12.2 | 349 | 11 | 057110 | 057110 | variola virus |
| 15 | 141.5 | 12.2 | 349 | 11 | 057111 | 057111 | variola virus |
| 16 | 141.5 | 12.2 | 349 | 11 | 089118 | 089118 | variola virus |
| 17 | 141.5 | 12.1 | 349 | 11 | 089098 | 089098 | variola virus |
| 18 | 141 | 12.1 | 355 | 11 | 085308 | 085308 | cowpox virus |
| 19 | 139 | 12.0 | 277 | 2 | 014866 | 014866 | homo sapien |
| 20 | 139 | 12.0 | 426 | 2 | 014865 | 014865 | homo sapien |
| 21 | 139 | 12.0 | 348 | 11 | 057112 | 057112 | variola virus |
| 22 | 139 | 12.0 | 348 | 11 | 085407 | 085407 | variola virus |
| 23 | 138.5 | 11.9 | 351 | 11 | 057117 | 057117 | cowpox virus |
| 24 | 137.5 | 11.8 | 351 | 11 | 057121 | 057121 | cowpox virus |
| 25 | 136.5 | 11.7 | 349 | 11 | 057284 | 057284 | camelipox virus |
| 26 | 136.5 | 11.7 | 349 | 11 | 057098 | 057098 | camelipox virus |
| 27 | 136 | 11.7 | 186 | 11 | 072735 | 072735 | cowpox virus |
| 28 | 135 | 11.6 | 349 | 11 | 057305 | 057305 | cowpox virus |
| 29 | 132.5 | 11.4 | 380 | 2 | 000280 | 000280 | homo sapien |

| | | | | | | | | |
|----|-------|------|-----|----|--------|--------|------------|---------|
| 30 | 133.5 | 11.4 | 350 | 11 | 057123 | 057123 | compox | virtu |
| 31 | 130.5 | 11.2 | 349 | 11 | 057097 | 057097 | camelop | virtu |
| 32 | 130.5 | 11.2 | 349 | 11 | 057109 | 057109 | variola | virtu |
| 33 | 127.5 | 11.0 | 326 | 11 | 057120 | 057120 | compox | virtu |
| 34 | 127.5 | 11.0 | 326 | 11 | 057122 | 057122 | compox | virtu |
| 35 | 126.5 | 10.9 | 347 | 11 | 057115 | 057115 | compox | virtu |
| 36 | 126.5 | 10.9 | 360 | 11 | 057118 | 057118 | compox | virtu |
| 37 | 126.5 | 10.9 | 347 | 11 | 057119 | 057119 | compox | virtu |
| 38 | 122 | 10.5 | 372 | 2 | 000279 | 000279 | homo | saplen |
| 39 | 121.5 | 10.5 | 418 | 2 | 000275 | 000275 | homo | saplen |
| 40 | 121.5 | 10.5 | 401 | 10 | 070202 | 070202 | mus | musculu |
| 41 | 121 | 10.4 | 425 | 2 | 000276 | 000276 | homo | saplen |
| 42 | 118.5 | 10.2 | 253 | 2 | 016042 | 016042 | homo | saplen |
| 43 | 116.5 | 10.0 | 401 | 10 | 008712 | 008712 | mus | musculu |
| 44 | 115.5 | 9.9 | 334 | 3 | 024403 | 024403 | trosophila | |
| 45 | 114.5 | 9.9 | 869 | 12 | 042126 | 042126 | xenopus | laeae |

ALIGNMENTS

| RESULT | 1 |
|--------|---|
| 057277 | |
| ID | 057277 |
| AC | 057277 |
| DT | 01-JUN-1998 (TREMBLREL, 06, PRELIMINARY; PERT; 348 AA. |
| DT | 01-JUN-1998 (TREMBLREL, 06, CREATED) |
| D1 | 01-JUN-1998 (TREMBLREL, 06, LAST SEQUENCE UPDATE) |
| D7 | 01-JUN-1998 (TREMBLREL, 06, LAST ANNOTATION UPDATE) |
| DE | TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG. |
| GN | CRMB. |
| OS | MONKEYPOX VIRUS. |
| OC | VIRUSES: DSNNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE. |
| CC | ORHPOFOXVIRUS. |
| RM | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN-ZAIRE-1996(96-17), ZAIRE-1996 (96-16); |
| RA | LOPAREV V.N., PARSONS J.M., ESPPOSITO J.J. ; |
| RL | SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBD3 DATA BANKS. |
| DR | EMBL: U88543; G2738105; - |
| DR | EMBL: U87841; G2738061; - |
| Q0 | PROSITE; PS00652; TNFR_NFR_1; 2. |
| Q0 | SEQUENCE 348 AA; 38212 MW; 1AAAB69 CRC32; |

| | | | | |
|-----------------------|-------|-----------------|--------|---------------|
| Query Match | 13.8% | Score 160.5 | DB 11 | Length 348 |
| Best Local Similarity | 23.9% | Pred. No. 2e-05 | | |
| Matches | 56 | Conservative | 28 | Mismatches 71 |
| | | | Indels | 79 |
| | | | Gaps | 12 |

| | | | | |
|----|-----|--|-----------------------------|--------------|
| QY | 10 | RTVFLAAILFLHLHLAKKSC----- | ENGDCRQEFKDRSGNCVLCQCGPGM | 56 |
| | | :: | | |
| | | :: | | |
| Db | 2 | RSVLYXSYLFL----- | SCIINRDLAPAPNSGKCKDNRYRSR-- | LCCLSGPPT 51 |
| | | | :: | |
| | | | :: | |
| QY | 57 | ELSKCGGCGYGDAOCVPCRPHRF-KEDMGFOKCPACADALVNRFORANGHTSDAVCG | | 115 |
| | | :: | :: | |
| | | :: | :: | |
| Db | 52 | YASRLC--DSKINQCTPCGSDPTFTSHNNHLQACLSGNCRCDNSOVETRSCTYTHNRIC- | | 109 |
| | | | :: | |
| | | | :: | |
| QY | 116 | DCLPEFY----- | RKTKL-VGF | |
| | | :: | | |
| | | :: | | |
| Db | 109 | ECSPEYICLLKSGSCCRICISKTKGIGIGVSGYSTGVIICSPCG--PGTYSHTYSSSTD | | 166 |
| | | | :: | |
| | | | :: | |
| QY | 146 | --EPHCTSKVNLVYKISSTVSSPRDTAAVAATGALATVLLACSCVYSFARGSS | | 197 |
| | | :: | | |
| | | :: | | |
| Db | 167 | KCEPTSTMTFNTDVEINLMPVNDI----- | SCRRTTGTGLS 202 | |
| | | | :: | |
| | | | :: | |

| | | |
|--------|--|---|
| RESULT | 2 | |
| 057108 | | |
| AC | 057108 | PRELIMINARY; PRT; 348 AA. |
| DT | 01-JUN-1998 | (TREMBLREL. 06, CREATED) |
| DT | 01-JUN-1998 | (TREMBLREL. 06, LAST SEQUENCE UPDATE) |
| DT | 01-AUG-1998 | (TREMBLREL. 07, LAST ANNOTATION UPDATE) |
| DE | TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG. | |

AC 092956;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE HERPESVIRUS ENTRY MEDIATOR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
CC CATARRHINI; HOMIINAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 97053782.
RA MONTGOMERY R.I., WARNER M.S., LDM B.J., SPEAR P.G.;
RT "Herpes simplex virus-1 entry into cells mediated by a novel member of
RL the TNF/NF receptor family";
RL CELL 87:427-436(1996).
DR EMBL: U70321; E319244;
DR PROSITE: PS00652; TNFR_NGFR_1;
DR PFAM: PF00020; TNFR_C6; 2.
SQ SEQUENCE 283 AA; 30420 MW; B67FBE7B CRC32;

Query Match 12.3%; Score 142.5; DB 2; Length 283;
Best Local Similarity 25.4%; Pred. No. 0.00039;
Matches 60; Conservative 28; Mismatches 83; Indels 65; Gaps 16;
QY 10 RTVFLALFLHLAKCYSCET-GDCRQDFKDRSGNCVLCQCGPMELSKCGFGYGE 68
DB 17 RTDVLRLVLYLFLAPCAPALPSCKEDEYVGS-EC--CPKCSPPGYKKEACGELTG- 73
QY 69 DAQCPKCPHREKEDW-GPKCKPCADC--ALVNRFORANCSTSDAVCGDCLPGFY--- 123
DB 73 -TVCEPCPPGTIAHLNGSKCLQCMCPAGLRASR-NCSTRTEAAYCG-CSPGHFCTIV 129
QY 123 -----RKTLYG--FQDMCEVPCGDPPEYEPHCT-----SKVN- 155
DB 130 QDGDHCAACRAVATSPGQGVKGGTESGDTLCQNC--PPGTFSPNGTLEECQHQTCSW 187
QY 155 -LVKISSTVSSRR-----DTAVALVICALATVLLACSSVSTARGSSRRNP 202
DB 188 LVTKAGAGTSSSHWVWVWFLSGSLVIVYVCSVTGLIIT---CVK-----RRKP 230

RESULT 13
035305 PRELIMINARY; PRT; 625 AA.
AC 035305;
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE RECEPTOR ACTIVATOR OF NF-KAPPA (RANK).
GN RANK.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
CC SCIROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RX TISSUE-LIVER EPITHELIUM;
RX MEDLINE: 98032977.
RA ANDERSON D.M., MARASKOVSKY E., BILLINGSLEY W.L., DOUGALL W.C.,
RA TOMETSKO M.E., ROUX E.R., TEEPE M.C., DUBOSE R.F., COSMAN D.,
RA GALBERT L.;
RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
RT and dendritic-cell function";
RL NATURE 390:175-179(1997).
DR EMBL: AF019046; G2612920;
DR MGD: MGI:1194917; RANK.
DR PFAM: PF00020; TNFR_C6; 3.
SQ SEQUENCE 625 AA; 66621 MW; A01E6C5B CRC32;

Query Match 12.2%; Score 142; DB 10; Length 625;
Best Local Similarity 27.0%; Pred. No. 0.00087;
Matches 54; Conservative 26; Mismatches 76; Indels 44; Gaps 12;

QY 13 LFAALFLHLAKCYSCETGDCRQDFKDRSGNCVLCQCGPMELSKCGFGYGEADAC 72
DB 15 LIALCVLLVPLQVLTQV--TPCTQERHXYHLGRC--CSRCEPGKYLSSKC--TPTSDDVC 69
QY 73 VPCRPKREKEDWGFQ-KC---KPC-ADALV-----NFRFORANCST-----SDAVC-- 115
DB 70 LFCGPDEYLDTWNEEDKCLHLKVCAGAKALVADPGNHTAPRCACCTAGYHNSDCECCR 129
QY 115 --GDCLPGEYRKTLYGFDMECVPC-----GDPPEYEPHCTSKVNLVYKIST 161
DB 130 RNTCEAPFPGAGNHPQLKMKDVTCTPCLLGFPSDFVSSYDKCKPFW-TNCT-----LLGK 181
QY 162 VSSPRLTAVALVICALATV 181
DB 182 LEAHGTTESDVCSSMTL 201

RESULT 14
057110 PRELIMINARY; PRT; 349 AA.
AC 057110;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS VARIOLA VIRUS.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYIRINAE; CHORDOPOXYIRINAE;
CC ORTHOPOXYVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-MAJOR, CONGO-1970 (CNG-70);
RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U88147; G2738093;
DR PROSITE: PS00652; TNFR_NGFR_1; 2.
SQ SEQUENCE 349 AA; 38273 MW; BEC101E8 CRC32;

Query Match 12.2%; Score 141.5; DB 11; Length 349;
Best Local Similarity 24.2%; Pred. No. 0.00056;
Matches 51; Conservative 30; Mismatches 87; Indels 43; Gaps 10;
QY 10 RTVFLALFLHL-----ACKVCSETGDCRQDFKDRSGNCVLCQCGPMELSKCG 63
DB 2 KSVLYLILFLSCILINGRDAPYTPPNKCKKDTYK-RHNC--CLSCPCTYASRLC- 58
QY 64 FGYGDAQCVPKPRHF-KEDMGFKCKPCADCALVNRFORANCSTSDAVCGDCLPGFY 122
DB 58 -DSKTNCTCPCSGSTFTSRNHLPLACLSNCRGNSNYETRSCVTTHNRIC-ECSPGY 115
QY 123 -----RKTLY-VGF-----QDMCEVPCGDPPEY-----EPHCSTK 152
DB 116 CLKGSSGCKACVSTQTKGIGYSGHYSVGDYISPCGFGTYSTVSADKCEVPNNT 175
QY 153 VNLVKISSTVSSPRLTAVALVICALATVIL 183
DB 176 FVITVEITLIVVNDTSCRTTITGLSESIL 206

RESULT 15
057111 PRELIMINARY; PRT; 349 AA.
AC 057111;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
GN CRMB.
OS VARIOLA VIRUS.
OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYIRINAE; CHORDOPOXYIRINAE;
CC ORTHOPOXYVIRUS.

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-MINOR, SIERRA LEONE-1968 (SLN-68);
RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: U88150; G2738099; -
DR PROSITE: PS00652; TNER_NGFR_1; 2.
SQ SEQUENCE 349 AA; 38209 MW; 8DAC3FB5 CRC32;

Query Match 12.2%; Score 141.5; DB 11; Length 349;
Best Local Similarity 24.2%; Pred. No. 0.00056;
Matches 51; Conservative 30; Mismatches 87; Indels 43; Gaps 10;

QY 10 RTVLFAILFLHL-----ACKVSCETGDCRQEFDRSGNCVLCKQGGPMGLSKECG 63
DB 2 KSVLYLXILFLSCIINGRDAAYTPPENGCKDTEYK-RHNLK-CLSCPPTGYASRLC- 58
QY 64 FGYGEDAQCVPKPRHPR-KEDWGFQKCKPCADCALVNRFORANCSTSDAVCGDCLPGFY 122
DB 58 -DSKTNTOCTPCGSGTFTSRNNHLPACLSGNGRONSQVETRSQNTNHRIC-ECSFGY 115
QY 123 -----RKTKL-VGF-----QDMCVPCGDDPPPY-----EPHCTSK 152
DB 116 CLKSSGCGACYSQYKCGIGYGVSGHTSVGDYICSPCGFGTYSTYVSTDKCEPVNNT 175
QY 153 VNLVKISSTVSPRDTAVAAVICSALATVLL 183
DB 176 FNYIDVEITLYPYNDISCTRTTGTGLSESTL 206

Search completed: May 14, 1999, 21:39:58
Job time: 7992 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 17:21:18 ; Search time 42.79 seconds

(without alignments)
71.372 Million cell updates/secTitle: US-09-212-270-4
Perfect score: 873

Sequence: 1 MALKVLPLHRTVLFPAILFL.....QDMECVPCGDPPEPPPEPHCE 151

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database: A_Geneseq_34:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|----------|--------------------|
| 1 | 714.5 | 81.8 | 417 | 1 W70386 | Amino acid sequenc |
| 2 | 714.5 | 81.8 | 423 | 1 W70387 | Amino acid sequenc |
| 3 | 135.5 | 13.5 | 438 | 1 R81882 | Plasmod PDC406/OX4 |
| 4 | 135.5 | 13.5 | 206 | 1 R81881 | Mouse type-II memb |
| 5 | 135.5 | 15.5 | 438 | 1 W48976 | OX40/Fc mutein. Pu |
| 6 | 135.5 | 15.5 | 206 | 1 W48977 | Mouse OX40 extrac |
| 7 | 129.5 | 14.8 | 833 | 1 W64466 | Human DR3 protein. |
| 8 | 129 | 14.8 | 625 | 1 W69958 | Murine NF-kB recep |
| 9 | 129 | 14.8 | 616 | 1 W69952 | NF-kB receptor act |
| 10 | 129 | 14.8 | 451 | 1 W69951 | RANK polypeptide p |
| 11 | 129 | 14.8 | 451 | 1 W68287 | RANK polypeptide p |
| 12 | 129 | 14.8 | 616 | 1 W68288 | NF-kB receptor act |
| 13 | 129 | 14.8 | 625 | 1 W68294 | Murine NF-kB recep |
| 14 | 128.5 | 14.7 | 283 | 1 W05809 | Human tumour necro |
| 15 | 128.5 | 14.7 | 283 | 1 W69238 | Herpesvirus entry |
| 16 | 128 | 14.7 | 616 | 1 W79233 | A receptor protein |
| 17 | 127.5 | 14.6 | 283 | 1 W12659 | Human herpes simpl |
| 18 | 127 | 14.5 | 466 | 1 R07447 | Human laminin B1 c |
| 19 | 126 | 14.4 | 186 | 1 R62655 | Cowpox virus psc 1 |
| 20 | 126 | 14.4 | 1786 | 1 W50893 | Human laminin B1 c |
| 21 | 125.5 | 14.4 | 283 | 1 W60045 | Human TNF receptor |
| 22 | 125 | 14.3 | 197 | 1 W23220 | Extracellular doma |
| 23 | 124.5 | 14.3 | 355 | 1 R85073 | Cowpox virus T2-eg |
| 24 | 123.5 | 14.1 | 417 | 1 W31517 | Death domain conta |
| 25 | 123.5 | 14.1 | 428 | 1 W31516 | Death domain conta |
| 26 | 123.5 | 14.1 | 181 | 1 W26708 | Human apoptosis pr |
| 27 | 123.5 | 14.1 | 417 | 1 W26709 | Human apoptosis pr |
| 28 | 123.5 | 14.1 | 417 | 1 W57045 | Human apoptosis in |
| 29 | 120 | 13.7 | 1764 | 1 P91672 | Primary amino acid |
| 30 | 120 | 13.7 | 1776 | 1 W50894 | Mouse laminin B1 c |
| 31 | 119 | 13.6 | 176 | 1 W80254 | Amino acid sequenc |
| 32 | 118 | 13.5 | 411 | 1 W57046 | Mouse apoptosis in |
| 33 | 117 | 13.4 | 277 | 1 R76996 | Deduced sequence e |
| 34 | 116 | 13.3 | 277 | 1 R74737 | ACT-4 cell surface |
| 35 | 116 | 13.3 | 277 | 1 R79904 | ACT-4-h-1 receptor |
| 36 | 115 | 13.2 | 326 | 1 R27866 | Myxoma virus T2 pr |
| 37 | 115 | 13.2 | 326 | 1 R85072 | Myxoma virus T2 pr |
| 38 | 114.5 | 13.1 | 277 | 1 R38859 | CD40 protein. CD40 |
| 39 | 111.5 | 12.8 | 591 | 1 W69950 | RANK polypeptide p |
| 40 | 111.5 | 12.8 | 591 | 1 W69950 | RANK polypeptide p |
| 41 | 110.5 | 12.7 | 1798 | 1 W50896 | Human laminin B2 c |
| 42 | 110 | 12.6 | 260 | 1 R91441 | Human CD27 antigen |
| 43 | 109.5 | 12.5 | 260 | 1 R20814 | T lymphocyte-speci |

44 108.5 12.4 255 1 R64197
45 108.5 12.4 255 1 R70977

Human 4-1BB polype
H4-1BB receptor pr

ALIGNMENTS

RESULT 1
ID W70386 standard; Protein; 417 AA.
AC W70386;
DT 02-DEC-1998 (first entry)
DE Amino acid sequence of human alpha-OARF065.
KW Human; alpha-OARF065; stroma cell; antibody; inflammatory;
OS cytokine-mediated disease; rheumatism; ulcerative colitis.
FH Key Location/Qualifiers
FT Misc-difference 223
FT Misc-difference 224 /note="encoded by AGA"
FT Misc-difference 224 /note="encoded by CCR"
FT W09838304-A1.
PD 03-SEP-1998.
PE 26-FEB-1998: J00799.
PR 27-FEB-1997: JP-043143.
PA (ONOX) ONO PHARM CO LTD.
PI Fukushima D, Konishi M, Tada H;
DR WPI: 98-481205/41.
DR N-PSDB: V33361.
PT Membrane polypeptide expressed by human stroma cells, and antibodies
PT recognising it - for treatment of inflammatory and other
PT cytokine-mediated diseases.
PS Claim 1: Pages 28-30; 54pp; Japanese.
CC This is the amino acid sequence of the human alpha-OARF065, used in
CC the method of the invention. The process involves the use of peptides
CC expressed by stroma cells, and its antibodies are used for in the
CC prevention and treatment of inflammatory and other cytokine-mediated
CC diseases such as rheumatism, ulcerative colitis.
SQ Sequence 417 AA;

Query Match 81.8%; Score 714.5; DB 1; Length 417;
Best Local Similarity 82.0%; Pred. No. 2.9e-57;
Matches 123; Conservative 9; Mismatches 17; Indels 1; Gaps 1;

QY 1 MALVLPPLHRTVLFPAILFLHLACKVSCETGDCSROEFDRSGNCVLCCKGPGMELS 60
DB 1 MALVLPPLHRTVLFPAILFLHLACKVSCETGDCSROEFDRSGNCVLCCKGPGMELS 59
QY 61 KECGFGYGEDAQVCPBPRHREKEDMGFOCKPCADCALVNRFORANCSHTSDAICGDCLP 120
DB 60 KECGFGYGEDAQVCPBPRHREKEDMGFOCKPCADCALVNRFORANCSHTSDAICGDCLP 119
QY 121 GFYRKTKLVGFQDMECVPCGDPPEPPPEPHC 150
DB 120 GFYRKTKLVGFQDMECVPCGDPPEPPPEPHC 149

RESULT 2
ID W70387 standard; Protein; 423 AA.
AC W70387;
DT 02-DEC-1998 (first entry)
DE Amino acid sequence of human beta-OARF065.
KW Human; beta-OARF065; stroma cell; antibody; inflammatory;
OS cytokine-mediated disease; rheumatism; ulcerative colitis.
FH Key Location/Qualifiers
FT Misc-difference 223
FT Misc-difference 224 /note="encoded by AGA"
FT Misc-difference 224 /note="encoded by CCR"
FT W09838304-A1.

KW RANKL; RANK ligand; tumour necrosis factor; TNF; murine.
 OS Mus sp.
 PN W09828426-A2.
 PD 02-JUL-1998.
 PE 22-DEC-1997; U23775.
 PR 14-OCT-1997; US-064671.
 PR 23-DEC-1996; US-059978.
 PR 07-MAR-1997; US-813509.
 PA (IMMUNEX CORP.
 PI Anderson DM, Galibert LJ, Maraskovsky E;
 DR WPI: 98-377657/32.
 DR N-PSDB: VA1375.
 PT New isolated ligand for receptor activator of NF-kappa B - used to
 PT develop products for augmenting an immune response for inhibiting an
 PT inflammatory response and for protection of cells
 PS Example 14; Pages 62-64; 80pp; English.
 CC This represents the murine mURANK (receptor activator of necrosis
 CC factor-kappaB (NF-kB)) polypeptide which is a homolog of the human RANK.
 CC RANK is a member of the tumour necrosis factor (TNF) family. A soluble
 CC expressing membrane-associated RANK with a soluble RANK which binds to
 CC RANK ligand (RANKL). RANKL polypeptides can activate RANK and can be
 CC used to induce maturation of dendritic cells and enhance their
 CC allo-stimulatory capacity, thereby augmenting an immune response. The
 CC soluble RANK polypeptide composition may also be used for regulating an
 CC immune or inflammatory response. Inhibition of NF-kB by RANK antagonists
 CC may be useful in ameliorating negative effects of an inflammatory
 CC response that result from triggering of RANK, e.g. in treating toxic
 CC shock or sepsis, graft-versus-host reactions, or acute inflammatory
 CC reactions. They can also be used in adjunct therapy for disease
 CC characterised by neoplastic cells that express RANK. RANKL polypeptides
 CC can also be used to identify inhibitors of RANK and thus inhibitors of
 CC an inflammatory response, and also for protecting RANK-expressing cells
 CC from the negative effects of chemotherapy or the presence of high levels
 CC of TNF-alpha. The products can also be used for detection and drug
 CC screening.
 SQ Sequence 625 AA;

Query Match 14.8%; Score 129; DB 1; Length 625;
 Best Local Similarity 29.9%; Pred. No. 0.00023;
 Matches 44; Conservative 20; Mismatches 57; Indels 26; Gaps 10;

QY 13 LPAALFLHLACKYSCETGDCSROQEFKDRSGNCVLCQSGPMELSKGCGYGEDAQ 72
 Db 15 LALCVLLVLPQVLTQV--TPCTGRHY-EHLGRC--CSRCPGRYLSKRC--TPTSQV 68
 QY 73 CVPCPRHFRKEDMGFO-KC---KPC-ADCALY-----NRFQANCSHT-----SDAVC- 116
 Db 69 CLPCGPDEYLDVTWNEEDKCLHKKVCDAGKALVADPDGNHTAPRCACAGYHMNSDCBCC 128
 QY 116 --GDCLPFFYRKTGLVGFQDMCEVCP 139
 Db 129 RRMTECARGFQAQHPQLQNKDVTCTPC 155

RESULT 9
 W69952
 ID W69952 standard; Protein; 616 AA.
 AC W69952;
 DT 08-OCT-1998 (first entry)
 DE NF-kB receptor activator RANK.
 KW RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
 KW immune response; inflammatory response; toxic shock; sepsis;
 KW RANKL; RANK ligand; tumour necrosis factor; TNF.
 OS Homo sapiens.
 PN W09828426-A2.
 PD 02-JUL-1998.
 PE 22-DEC-1997; U23775.
 PR 14-OCT-1997; US-064671.
 PR 23-DEC-1996; US-059978.
 PR 07-MAR-1997; US-813509.
 PA (IMMUNEX CORP.
 PI Anderson DM, Galibert LJ, Maraskovsky E;
 DR WPI: 98-377657/32.
 DR N-PSDB: VA1375.
 PT New isolated ligand for receptor activator of NF-kappa B - used to
 PT develop products for augmenting an immune response for inhibiting an
 PT inflammatory response and for protection of cells

PI Anderson DM, Galibert LJ, Maraskovsky E;
 DR WPI: 98-377657/32.
 DR N-PSDB: VA1376.
 PT New isolated ligand for receptor activator of NF-kappa B - used to
 PT develop products for augmenting an immune response for inhibiting an
 PT inflammatory response and for protection of cells
 PS Example 1; Pages 48-51; 80pp; English.
 CC This represents the polypeptide RANK (receptor activator of necrosis
 CC factor-kappaB (NF-kB)). RANK is a member of the tumour necrosis factor
 CC (TNF) family. A soluble RANK may be used for inhibiting activation of
 CC NF-kB, by contacting a cell expressing membrane-associated RANK with a
 CC soluble RANK which binds to RANK ligand (RANKL). RANKL polypeptides can
 CC activate RANK and can be used to induce maturation of dendritic cells and
 CC enhance their allo-stimulatory capacity, thereby augmenting an immune
 CC response. The soluble RANK polypeptide composition may also be used for
 CC regulating an immune or inflammatory response. Inhibition of NF-kB by
 CC RANK antagonists may be useful in ameliorating negative effects of an
 CC inflammatory response that result from triggering of RANK, e.g. in
 CC treating toxic shock or sepsis, graft-versus-host reactions, or acute
 CC inflammatory reactions. They can also be used in adjunct therapy for
 CC disease characterised by neoplastic cells that express RANK. RANKL
 CC polypeptides can also be used to identify inhibitors of RANK and thus
 CC inhibitors of an inflammatory response, and also for protecting RANK-
 CC expressing cells from the negative effects of chemotherapy or the
 CC presence of high levels of TNF-alpha. The products can also be used for
 CC detection and drug screening.
 SQ Sequence 616 AA;

Query Match 14.8%; Score 129; DB 1; Length 616;
 Best Local Similarity 21.5%; Pred. No. 0.00023;
 Matches 45; Conservative 28; Mismatches 60; Indels 76; Gaps 12;

QY 7 PLHRVFLAALFLHLACKYSCETGDCSROQEFKDRSGNCVLCQSGPMELSKGCGFG 66
 Db 10 PLFALLLCALARLQVALQIA--PCTSEKH-EHLGRC--CNKCPGRYMSKRC--T 61
 QY 67 XGEDACVPCRPFRKEDMGFO-KC-----KPCA----- 95
 Db 62 TTSDSVCLPCGPDYLDWSMNEEDKCLHKKVCDTGKALVAVAGNSTPTRCAGTAGYHWS 121
 QY 95 -DCALVNRFORANCS-----H---TSDAVCGDCLPGFY----- 124
 Db 122 QDCECRR--NTECARGAQAHPQLQNKDVTCTPCRLAIFYDAFSSDTCKRPTWCTFLG 179
 QY 124 RKTGLVGFQDMCEVCPGDP--PPEBPH 149
 Db 180 KRVEHGTGKSDAVCSSLPARKPNEPH 208

RESULT 10
 W69951
 ID W69951 standard; Protein; 451 AA.
 AC W69951;
 DT 08-OCT-1998 (first entry)
 DE RANK polypeptide partial sequence encoded by cDNA clone 9D-15C.
 KW RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
 KW immune response; inflammatory response; toxic shock; sepsis;
 KW RANKL; RANK ligand; tumour necrosis factor; TNF.
 OS Homo sapiens.
 PN W09828426-A2.
 PD 02-JUL-1998.
 PE 22-DEC-1997; U23775.
 PR 14-OCT-1997; US-064671.
 PR 23-DEC-1996; US-059978.
 PR 07-MAR-1997; US-813509.
 PA (IMMUNEX CORP.
 PI Anderson DM, Galibert LJ, Maraskovsky E;
 DR WPI: 98-377657/32.
 DR N-PSDB: VA1375.
 PT New isolated ligand for receptor activator of NF-kappa B - used to
 PT develop products for augmenting an immune response for inhibiting an
 PT inflammatory response and for protection of cells

PS Example1: Pages 44-46: 80pp. English
CC This represents a partial sequence of the polypeptide RANK (receptor
CC activator of necrosis factor- κ appb (NF- κ B)). RANK is a member of
CC the tumour necrosis factor family. A soluble RANK may be used for
CC inhibiting activation of NF- κ B, by contacting a cell expressing membrane
CC -associated RANK with a soluble RANK which binds to RANK ligand (RANKL).
CC RANKL polypeptides can activate RANK and can be used to induce maturation
CC of dendritic cells and enhance their allo-stimulatory capacity, thereby
CC augmenting an immune response. The soluble RANK polypeptide composition
CC may also be used for regulating an immune or inflammatory response.
CC Inhibition of NF- κ B by RANK antagonists may be useful in ameliorating
CC negative effects of an inflammatory response that result from triggering
CC of RANK, e.g. in treating toxic shock or sepsis, graft-versus-host
CC reactions, or acute inflammatory reactions. They can also be used in
CC adjunct therapy for disease characterised by neoplastic cells that
CC express RANK. RANKL polypeptides can also be used to identify inhibitors
CC of RANK and thus inhibitors of an inflammatory response, and also for
CC protecting RANK-expressing cells from the negative effects of
CC chemotherapy or for the presence of high levels of TNF- α . The products
CC can also be used for detection and drug screening.
50 Sequence 451 AA;

| | | | | |
|-----------------------|-----------------|-------------------|-----------|------------|
| Query Match | 14.8% | Score 129 | DB 1 | Length 451 |
| Best Local Similarity | 21.5% | Pred. No. 0.00017 | | |
| Matches 45 | Conservative 28 | Mismatches 60 | Indels 76 | Gaps 12 |

QY 7 PLHRTVFAAILFLLHLACKVSCETGDCSRQQEFKDRSGNCVLCCKQCGPMELSKCEGFG 66

Db 10 PLFALLLCALRLQVALQIA--PPCTSEKHY-EHLGRC--CNKCEPGKYMSSKC--T 61

QY 67 YGEDAQCVP C R P H R F K E D W G F Q - K C - - - - - K P C A - - - - - 95

Db 62 TTSDSVCLPCGPDEYLD SWNEEDKCLHKVCDTGKALVAVVAGNSTTPRCAC TAGYHWS 121

| | | | |
|----|----|---|-----|
| QY | 95 | -DCALVNRQRANC-----H-----TSDAVCGDCLPGFY----- | 124 |
|----|----|---|-----|

Db 122 QDCECCRR--NTECAPGLGAQHPLQLNKDVTVCCKPCLAGYFSDAFSSTDCKRPWTNCTFLG 179

QY 124 RKT^{KL}VG^{FQ}DM^{EC}VP^{CG}DP⁻⁻⁻PRYEPH 149

Db 180 KVEHHGTEKSDAVCSSSLPARKPPNEPH 208

| | |
|--------|----------------------------------|
| RESULT | 11 |
| W68287 | |
| ID | W68287 standard; Protein; 451 AA |

DE 08-OCT-1998 (first entry)
DR RANK polypeptide partial sequence encoded by cDNA clone 9D-15C.
KW RANK: necrosis factor-kappa B; NF-kB; receptor activator; human;
KW immune response; inflammatory response; toxic shock; sepsis;
KW RANKL; RANK ligand; tumour necrosis factor; TNF.
OS Homo sapiens.
PN M09828424-A2.
PD 02-JUL-1998.
PE 22-DEC-1997; U23866.
PR 14-OCT-1997; US-064671.
PR 23-DEC-1996; US-059978.
PR 07-MAR-1997; US-813509.
PA (IMV) IMMUNEX CORP.
PI Anderson DM, Galibert LJ, Maraskovsky E;
DR WPI; 98-377655/32.
DR N-PSDB; V41369.
PT New isolated receptor activator of necrosis factor-kappa B - useful
PT for e.g. developing products for regulating an immune or
PT inflammatory response, treating toxic shock or sepsis
PS Example 1; Pages 44-46; 80pp; English.
CC This represents a partial sequence of the polypeptide RANK (receptor
CC activator of necrosis factor-kappaB (NF-kB)). RANK is a member of
CC the tumour necrosis factor family. Host cells transformed or transfected
CC with an expression vector comprising the RANK encoding nucleic acid can
CC be used to produce recombinant RANK protein. The soluble RANK may be used

for inhibiting activation of NF- κ B, by contacting a cell expressing membrane-associated RANK with a soluble RANK which binds to RANK ligand (RANKL). The soluble RANK polypeptide composition may also be used for regulating an immune or inflammatory response. Inhibition of NF- κ B by RANK antagonists may be useful in ameliorating negative effects of an inflammatory response that result from triggering of RANK, e.g. in treating toxic shock or sepsis, graft-versus-host reactions, or acute inflammatory reactions. They can also be used in adjunct therapy for disease characterised by neoplastic cells that express RANK. The products can also be used for detection and drug screening.

| | | | | |
|-----------------------|-----------------|-------------------|-----------|------------|
| Query Match | 14.8% | Score 129 | DB 1 | Length 451 |
| Best Local Similarity | 21.58% | Pred. No. 0.00017 | | |
| Matches 45 | Conservative 28 | Mismatches 60 | Indels 76 | Gaps 12 |

QY 7 PLHRTVLEAILELLHLACKVSCETGDCSRQQEFKDRSGNCVLCKQCGPGMELSKECGFG 66

Db 10 PLFALLLCALLQVALQIA--PCTSEKHY-EHLGRC--CNKCEPGKYMSSKC--T 61

| | | | |
|----|----|--|----|
| QY | 67 | YGEDAQCPCHPHRFKEDWGFQ-KC-----KPCA----- | 95 |
|----|----|--|----|

Db 62 TTSDSVCLPCGPDEYLD SWNEEDKCLHKVCDTGKALVAVVAGNSTTPRCAC TAGYHWS 121

| | | |
|----|---|-----|
| QY | 95 -DCALVNRQRANCS-----H-----TSDAVCGDCLPGFY----- | 124 |
|----|---|-----|

Db 122 QDCECCR--NTECAPGLGAQHPLQLNKDVTCKPCLAGYFSDAFSSSTDCKRPWTNCTFLG 179

QY 124 RKT^{KL}VGFDMECVPCGDP--PRYEPH 149

Db 180 KVEHHGTEKSDAVCSSSLPARKPNEPH 208

| | |
|--------|----------------------------------|
| RESULT | 12 |
| W68288 | |
| ID | W68288 standard; Protein; 616 AA |

08-OCT-1998 (first entry)
DE NF-kB receptor activator RANK.
DI RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
KM immune response; inflammatory response; toxic shock; sepsis;
KM RANKL; RANK ligand; tumour necrosis factor; TNF.
OS Homo sapiens.
PN WO9828424-A2.
PD 02-JUL-1998.
PE 22-DEC-1997; U23866.
PR 14-OCT-1997; US-064671.
PR 23-DEC-1996; US-059378.
PR 07-MAR-1997; US-813509.
PI (IMMV) IMMUNEX CORP.
PI Anderson DM, Galibert LJ, Maraskovsky E;
PI WPI: 98-377655/32.
DR N-PSDB: V41370
PT New isolated receptor activator of necrosis factor-kappa B - useful
PT for, e.g. developing products for regulating an immune or
PT inflammatory response, treating toxic shock or sepsis
PS Claim 27; Pages 48-51; 80pp; English.
CC This represents the polypeptide RANK (receptor activator of necrosis
CC factor-kappaB (NF-kB)). RANK is a member of the tumour necrosis factor
CC (TNF) family. Host cells transformed or transfected with an expression
CC vector comprising the RANK encoding nucleic acid can be used to produce
CC recombinant RANK protein. The soluble RANK may be used for inhibiting
CC activation of NF-kB, by contacting a cell expressing membrane-associated
CC RANK with a soluble RANK which binds to RANK ligand (RANKL). The soluble
CC RANK polypeptide composition may also be used for regulating an immune
CC or inflammatory response. Inhibition of NF-kB by RANK antagonists may be
CC useful in ameliorating negative effects of an inflammatory response that
CC result from triggering of RANK, e.g. in treating toxic shock or sepsis,
CC graft-versus-host reactions, or acute inflammatory reactions. They can
CC also be used in adjunct therapy for disease characterised by neoplastic
CC cells that express RANK. The products can also be used for detection and
CC drug screening.

OY 124 -----RKTIVG--FQDMCVPCGDPPEPEH 149
DB 130 ODGDHCACRAATATSSPGORVOKGTESODTLQNC--PPGTFFSPN 173

RESULT 15

W69238
ID W69238 standard; Protein: 283 AA.
AC W69238.
DT 21-Oct-1998 (first entry)
DE Herpesvirus entry mediator protein.
KW Entry mediator gene; herpesvirus; HVEM; tumour necrosis factor receptor;
KW gene expression regulator; cellular stress; inflammatory response;
KW lymphocyte activity regulator; autoimmune response.
OS Homo sapiens.
FN W09825967-A1.
PD 18-JUN-1998.
PF 05-DEC-1997; U22278.
PR 12-DEC-1996; US-032705.
PA (GETH) GENENTECH INC.
PI Ashkenazi AJ, Marsters SA;
DR WPI: 98-348457/30.
N-PSDB: V44852.
PT Herpesvirus entry mediator polypeptide, HVEM - useful, e.g. in
PT assays for HVEM and to produce antibodies and transgenic animals,
PT e.g. for drug screening
PS Claim 1: Fig 1: 46pp; English.
CC This sequence is the herpesvirus entry mediator (HVEM) protein of
CC the invention. The protein is useful in quantitative diagnostic assays
CC for HVEM, in affinity purification of HVEM from recombinant cells/natural
CC sources and in competitive-type receptor binding assays. It can also be
CC used to generate antibodies, also useful in diagnostic assays for HVEM
CC and affinity purification of HVEM. HVEM is believed to be a member of the
CC tumour necrosis factor receptor (TNFR) family, and transient transfection
CC of HVEM into human 293 cells caused marked activation of certain
CC transcription factors, e.g. AP-1, suggesting that HVEM is involved in
CC regulating gene expression in response to infectious stimuli and cellular
CC stress. The predominant expression of HVEM mRNA in lymphocyte-rich
CC tissues (e.g. spleen and peripheral blood) also suggested it may be a
CC receptor in regulating lymphocyte activity. Antibodies produced may
CC therefore be useful therapeutically, e.g. antagonistic antibodies may be
CC useful to block excessive inflammatory/autoimmune response resulting from
CC e.g. AP-1 induction. Whilst agonistic antibodies may enhance HVEM
CC regulation of such induction. The DNA may be used diagnostically, e.g.
CC to determine if DNA and/or RNA encoding HVEM is present in cells, and to
CC prepare HVEM polypeptide recombinantly. It is also useful to produce
CC non-human transgenic animals (e.g. mice or rats), especially knockout
CC animals containing cells with an altered gene encoding HVEM polypeptide.
CC Such animals are useful in the development and screening of
CC therapeutically useful reagents.
SQ Sequence 283 AA;

Query Match 14.7%; Score 128.5; DB 1: Length 283;
Best Local Similarity 27.7%; Pred. No. 0.00012;
Matches 46; Conservative 18; Mismatches 67; Indels 35; Gaps 10;

OY 10 RTVLAAILFLHLACKVSCETGDSRQOEFRDSGNCVLCQCGPGLSKECGFGYE 69
DB 17 RTDVRLVLYLFLGAPCAPALPSCKEDEYVGS--EC--CPKCSGYRVKKEAGELTG- 73
OY 70 DAOCVPCRPFRKEDW--GQCKCKPCADC--ALVNRFORANCSTSDAVCGDCLPGFY-- 124
DB 73 -TVCEPCPGIYIAHLNGSLKLCQCMCDPAAGLRASR--NCSRTENAVCG--CSPGHFCIV 129
OY 124 -----RKTIVG--FQDMCVPCGDPPEPEH 149
DB 130 ODGDHCACRAATATSSPGORVOKGTESODTLQNC--PPGTFFSPN 173

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 20:59:04 ; Search time 33.52 Seconds

(without alignments)
168.750 Million cell updates/sec

Title: US-09-212-270-4

Perfect score: 873
Sequence: 1 MALKVPLHRTVLFALILFL.....QDMCEVPCGDPPEPPPEHCE 151

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database: PIR_58:*

1: PIR1:*
2: PIR2:*
3: PIR3:*
4: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1 | 146.5 | 16.8 | 271 | 2 | S12783 | OX40 antigen precu |
| 2 | 135.5 | 15.5 | 272 | 2 | I48700 | gene ox40 protein |
| 3 | 126 | 14.4 | 1786 | 1 | MMHUB1 | laminin beta-1 cha |
| 4 | 125 | 14.3 | 435 | 2 | T54182 | tumor necrosis fac |
| 5 | 120 | 13.7 | 1786 | 1 | MMMSB1 | laminin beta-1 cha |
| 6 | 118.5 | 13.6 | 277 | 2 | A60771 | B-cell activation |
| 7 | 117 | 13.4 | 425 | 2 | A26431 | neve growth facto |
| 8 | 117 | 13.4 | 349 | 2 | D36858 | gene G4R protein - |
| 9 | 117 | 13.4 | 334 | 2 | A48151 | sperm tail protein |
| 10 | 116 | 13.3 | 277 | 2 | I37552 | OX40 homologue - h |
| 11 | 113 | 13.2 | 326 | 1 | GOVZML | T2 protein - myxom |
| 12 | 114 | 13.1 | 250 | 1 | A49053 | CD27 antigen precu |
| 13 | 111.5 | 12.8 | 427 | 1 | GOHUN | neve growth facto |
| 14 | 111 | 12.7 | 289 | 2 | A46515 | B cell associated |
| 15 | 111 | 12.7 | 305 | 2 | A46476 | CD40 - mouse |
| 16 | 110.5 | 12.7 | 260 | 1 | A46517 | CD27 antigen precu |
| 17 | 109 | 12.5 | 1792 | 1 | GOMST1 | tumor necrosis fac |
| 18 | 109 | 12.5 | 454 | 2 | I57826 | tumor necrosis fac |
| 19 | 109 | 12.5 | 265 | 2 | B48151 | sperm tail protein |
| 20 | 108.5 | 12.4 | 1607 | 1 | MMMSB2 | laminin gamma-1 ch |
| 21 | 108.5 | 12.4 | 255 | 2 | I38426 | 4-1BB - human |
| 22 | 108.5 | 12.4 | 255 | 2 | JT0752 | lymphocyte activat |
| 23 | 107.5 | 12.3 | 3084 | 1 | MMMSA | laminin alpha-1 ch |
| 24 | 106.5 | 12.2 | 1801 | 1 | MMRS | laminin beta-2 cha |
| 25 | 106.5 | 12.2 | 3712 | 2 | S18253 | variant-specific s |
| 26 | 104.5 | 12.0 | 596 | 2 | A45664 | gene murine tumour |
| 27 | 104 | 11.9 | 459 | 2 | I48854 | neve growth facto |
| 28 | 103.5 | 11.9 | 416 | 2 | JN0006 | laminin beta-2 cha |
| 29 | 103.5 | 11.9 | 1797 | 2 | A55677 | laminin beta-2 cha |
| 30 | 103.5 | 11.9 | 2482 | 2 | A47648 | perlecan homolog u |
| 31 | 103.5 | 11.9 | 2277 | 2 | B47648 | perlecan homolog u |
| 32 | 103.5 | 11.9 | 1160 | 2 | C47648 | perlecan homolog u |
| 33 | 103 | 11.8 | 1448 | 2 | B38634 | tumor necrosis fac |
| 34 | 102.5 | 11.7 | 1548 | 2 | S34583 | serine protease |
| 35 | 101.5 | 11.6 | 869 | 1 | MMHUB2 | laminin gamma-1 ch |
| 36 | 101 | 11.6 | 869 | 1 | JC4858 | VLDL receptor prec |
| 37 | 101 | 11.6 | 461 | 2 | A35356 | tumor necrosis fac |
| 38 | 101 | 11.6 | 3075 | 2 | S14458 | laminin alpha-1 ch |

| | | | | | | |
|----|-------|------|------|---|--------|--------------------|
| 40 | 100.5 | 11.5 | 1816 | 1 | S68960 | laminin alpha-4 ch |
| 41 | 100.5 | 11.5 | 1394 | 2 | A35626 | transforming growt |
| 42 | 99.5 | 11.4 | 873 | 1 | QRRBVD | VLDL receptor prec |
| 43 | 99.5 | 11.4 | 461 | 1 | GQRRT1 | tumor necrosis fac |
| 44 | 99 | 11.3 | 256 | 2 | B32393 | T-cell antigen 4-1 |
| 45 | 98.5 | 11.3 | 873 | 1 | A49729 | VLDL receptor prec |

ALIGNMENTS

RESULT 1
S12783
OX40 antigen precursor - rat
N:Alternate names: nerve growth factor receptor homolog
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 10-Sep-1997
C:Accession: S12783; S08036
R:Allet, S.; Fossum, S.; Barclay, A.N.
EMBO J. 9, 1063-1068, 1990
A>Title: Characterization of the MRC OX40 antigen of activated CD4 positive T lymphoc
A:Reference number: S12783; MUID:90214614
A:Accession: S12783
A:Molecule type: mRNA
A:Residues: 1-271 <MAL>
A:Cross-references: EMBL:X17037; NID:g57830; PID:g57831
C:Keywords: growth factor receptor; transmembrane protein
F:1-19/Domain: signal sequence #status predicted <Sig>
F:20-271/Product: OX40 antigen #status predicted <Mat>
F:211-235/Domain: transmembrane #status predicted <TM>

Query Match 16.8%; Score 146.5; DB 2; Length 271;
Best Local Similarity 30.1%; Pred. No. 4.9e-06;
Matches 34; Conservative 14; Mismatches 50; Indels 15; Gaps 5;
QY 46 NCY-----LCKOCGPMELSKECGFGYGDACVPCRPFRFEDMGCFCKPCADC 96
DB 25 NCVKDYTPSGHKCRCPGHWVSR--DHTDVTCHPCPEPFYNEAVNYDCKCCTQC 82
QY 97 ALVNRFO-RANGSHTSDAVCGDCLPGFYRTKLVGSDMEVCVCGDPPPEYEP 148
DB 83 NHRSGSELKQNCPTEDTVC-QCRPTQPRQDSSHLKGVDCVPC--PPGHFSP 132
RESULT 2
I48700
gene ox40 protein - mouse
N:Alternate names: OX40 antigen
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 18-Oct-1996
C:Accession: I48700; I48334; S34377
R:Calderhead, D.M.; Buhlmann, J.E.; van den Eertwegh, A.J.; Claassen, E.; Noelle, R.J
J. Immunol. 151, 5261-5271, 1993
A>Title: Cloning of mouse OX40: a T cell activation marker that may mediate T-B cell
A:Reference number: I48700; MUID:94044750
A:Accession: I48700
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-272 <RES>
A:Cross-references: EMBL:Z21674; NID:g312827; PID:g312828
R:Stielerland, M.L.; Copeland, N.G.; Gilbert, D.J.; Jenkins, N.A.; Barclay, A.N.
Eur. J. Immunol. 25, 926-930, 1995
A>Title: Gene structure and chromosomal localization of the mouse homologue of rat OX
A:Reference number: I48334; MUID:9555413
A:Accession: I48334
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-14, 'G', 16-272 <RE2>
A:Cross-references: EMBL:X85214; NID:g732818; PID:g732819
C:Genetics: OX40
A:Gene: OX40
A:Introns: 45/1; 86/1; 122/1; 144/2; 210/1; 250/1

Db 155 ELKDEVGKGNHCVCKAGHFQNTSSPSARCOPIHRC 192

RESULT 5

MMMSB1

laminin beta-1 chain precursor - mouse

N:Alternate names: laminin chain B1

C:Species: Mus musculus (house mouse)

C>Date: 28-Feb-1986 #sequence_revision 30-Jun-1991 #text_change 21-Aug-1998

C:Accession: A26413; S02679; S05326; S14877; A02871; S02036; S13543

R:Sasaki, M.; Kato, S.; Kohno, K.; Martin, G.R.; Yamada, Y.

Proc. Natl. Acad. Sci. U.S.A. 84, 935-939, 1987

A:Title: Sequence of the cDNA encoding the laminin B1 chain reveals a multidomain protein

A:Reference number: A26413; MUID:87147212

A:Accession: A26413

A:Molecule type: mRNA

A:Residues: 1-1786 <SAS>

A:Cross-references: EMBL:M15525; NID:q198700

A>Note: translation in Genbank has additional 48 residues at the amino end

R:Fujikawa, S.; Shinkai, H.; Deutzmann, R.; Paulsson, M.; Timpl, R.

Biochem. J. 252, 453-461, 1988

A:Title: Structure and distribution of N-linked oligosaccharide chains on various domain

A:Reference number: S02678; MUID:88326259

A:Accession: S02679

A:Molecule type: Protein

A:Residues: 28-42,932-946 <FUJ>

R:Hartl, L.; Oberdaumer, I.; Deutzmann, R.

Eur. J. Biochem. 173, 629-635, 1988

A:Title: The N terminus of laminin A chain is homologous to the B chains.

A:Reference number: S00624; MUID:88225080

A:Accession: S05326

A:Molecule type: Protein

A:Residues: 457-466;854-868;932-946 <HAR>

R:Mann, K.; Deutzmann, R.; Timpl, R.

Eur. J. Biochem. 178, 71-80, 1988

A:Title: Characterization of proteolytic fragments of the laminin-110 complex and th

A:Reference number: S08895; MUID:89078415

A:Accession: S14877

A:Molecule type: protein

A:Residues: 590-620 <MAN>

R:Barlow, D.P.; Green, N.M.; Kurkinen, M.; Hogan, B.L.M.

EMBO J. 3, 2355-2362, 1984

A:Title: Sequencing of laminin B chain cDNAs reveals C-terminal regions of coiled-coil

A:Reference number: A02870; MUID:85051302

A:Accession: A02871

A:Molecule type: mRNA

A:Residues: 1292-1530, 'MEMP', 1535-1691, 'C', 1693-1748, 'N', 1750-1786 <BAR>

A:Cross-references: EMBL:X05212; NID:952861; PID:g809042

R:Deutzmann, R.; Huber, J.; Schmetz, K.A.; Oberdaumer, I.; Hartl, L.

Eur. J. Biochem. 177, 35-45, 1988

A:Title: Structural study of long arm fragments of laminin. Evidence for repetitive C-te

A:Reference number: S01790; MUID:89030693

A:Accession: S02036

A:Molecule type: protein

A:Residues: 1561-1587 <DEU>

R:Paulsson, M.; Deutzmann, R.; Timpl, R.; Dalzoppo, D.; Odermatt, E.; Engel, J.

EMBO J. 4, 309-316, 1985

A:Title: Evidence for coiled-coil alpha-helical regions in the long arm of laminin.

A:Reference number: S13543; MUID:85257455

A:Accession: S13543

A:Molecule type: protein

A:Residues: 1700-1748, 'N', 1750-1759 <PAU>

C:Genetics:

A:Gene: Lamb-1

A:Map position: 12

C:Complex: Laminins are trimers of an alpha-type, a beta-type, and a gamma-type laminin

C:Function:

A:Description: Interact with cells and with other basement membrane proteins to promote

C:Superfamily: laminin beta-1 chain; laminin-type EGF-like homology

C:Keywords: basement membrane; calcium binding; cell binding; coiled coil; extracellular

F:1-21/Domain: signal sequence #status predicted <SIG>

F:22-1786/Product: laminin beta-1 chain #status predicted <MAT>

F:22-270/Domain: VI <DOM6>

F:271-540/Domain: V <DOM5>

F:271-332/Domain: laminin-type EGF-like homology <EG1>

F:335-395/Domain: laminin-type EGF-like homology <EG2>

F:398-455/Domain: laminin-type EGF-like homology <EG3>

F:458-507/Domain: laminin-type EGF-like homology <EG4>

F:510-540/Domain: laminin-type EGF-like homology #status atypical <EG5>

F:541-772/Domain: IV <DOM4>

F:773-1182/Domain: III <DOM3>

F:773-818/Domain: laminin-type EGF-like homology <EG6>

F:821-864/Domain: laminin-type EGF-like homology <EG7>

F:867-914/Domain: laminin-type EGF-like homology <EG8>

F:917-973/Domain: laminin-type EGF-like homology <EG9>

F:976-1025/Domain: laminin-type EGF-like homology <EG10>

F:1028-1081/Domain: laminin-type EGF-like homology <EG11>

F:1084-1129/Domain: laminin-type EGF-like homology <EG12>

F:1132-1176/Domain: laminin-type EGF-like homology <EG13>

F:1183-1397/Domain: II, heptad repeats <DOM2>

F:1398-1430/Domain: alpha <ALP>

F:1431-1786/Domain: I, heptad repeats <DOM1>

F:32/Modified site: pyrrolidone carboxylic acid (Gln) (in mature form) #status predic

F:120,356,519,677,1041,1195,1279,1336,1343,1487,1533,1542,1643/Binding site: carboxyhd

F:119,1182,1785/Disulfide bonds: interchain #status predicted

Query Match 13.7% Score 120; DB 1; Length 1786;

Best Local Similarity 24.5% Pred. No. 0.0039;

Matches 45; Conservative 18; Mismatches 57; Indels 64; Gaps 11;

QY 10 RTVFLALFLILH---LACKYS-----CET---GDCSRQDERRDRGNCVYLKQCCPG 56

Db 756 RNIIF-SISALIHQTGLACEDCDPQGSLSVCDPNQGCQCPNVVGR-----CNRCAFG 809

QY 57 MELSKECGFYGDACQVPCRPHR-----FREDGEF 87

Db 810 T-----FGEFPGN-CRPFCHLQGSASAFCDATGQCCHCFQGIYARQCRCRLPGYGF 861

QY 88 QKCRPC-----ADCAVNRFORANCSTSDAVGCDLPGFYRKTLVGFQDMCEVPCGD 141

Db 862 PSCGQCQNGHALDDYVTGSCLSQGDYTTGHNCRCLAGYV-GPPIIGSGD-HCRPCPC 919

QY 142 PPPP 145

Db 920 PDGP 923

RESULT 6

A60771

B-cell activation protein CD40 precursor - human

N:Alternate names: B-cell surface antigen Bp50

C:Species: Homo sapiens (man)

C>Date: 03-Jun-1993 #sequence_revision 03-Feb-1994 #text_change 10-Sep-1997

A:Accession: S04460; A60771

R:Stamenkovic, I.; Clark, E.A.; Seed, B.

EMBO J. 8, 1403-1410, 1989

A:Title: A B-lymphocyte activation molecule related to the nerve growth factor recept

A:Reference number: S04460; MUID:89356608

A:Accession: S04460

A:Molecule type: mRNA

A:Residues: 1-277 <STA>

A:Cross-references: EMBL:X60592; NID:929850; PID:q29851

R:Braesch-Andersen, S.; Paulie, S.; Koho, H.; Nika, H.; Aspenstroem, P.; Perlmann, P.

J. Immunol. 142, 562-567, 1989

A:Title: Biochemical characteristics and partial amino acid sequence of the receptor-

A:Reference number: A60771

A:Accession: A60771

A:Molecule type: protein

A:Residues: 21-50 <BRA>

A:Experimental source: Burkitt lymphoma cell line Raji

C:Genetics:

A:Gene: GDB:CD40

A:Cross-references: GDB:215268; OMIM:109535

A:Map position: 20q12-20q13.2


```
RESULT 9
A48151
Sperm tail protein Mst98ca - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 21-Jan-1994 #sequence_revision 18-Nov-1994 #text_change 24-Sep-1998
C:Accession: A48151; MUID:93180818
R:Schaefer, M.; Borsch, D.; Hulster, A.; Schaefer, U.
MOL Cell. Biol. 13, 1708-1718, 1993
A:Title: Expression of a gene duplication encoding conserved sperm tail proteins in tran
A:Reference number: A48151; MUID:93180818
A:Accession: A48151
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1534 <SCH>
A:Cross-references: EMBL:X67704, NID:98267, PID:98268
A:Note: sequence extracted from NCBI backbone (NCBIP:125984)
C:Genetics:
A:Gene: Mst98ca
A:Cross-references: FlyBase:FBgn0002865

Query Match 13.4%; Score 117; DB 2; Length 334;
Best Local Similarity 29.4%; Pred. No. 0.0021;
Matches 40; Conservative 5; Mismatches 47; Indels 44; Gaps 9;

QY 34 CSROGEFDRSGNCVLCQCGFGM-ELSKEG-FGVEDACVPCRPFRKEDWGFQKCK 91
DB 191 CFSTQVPAVEGRGCGPCGCGPCGRCPCGPG--PCGPGP-----RCG 237

QY 92 PCADCALVN-----RQRANCSHTSDAVCGDCLPGFYRTKTLVGFQDMECVPCG 140
DB 238 PGGPFAVNCGCGCGTLMTSGFPVPAFCPC--APCGPCP-----LCNSPCGPGC 285

QY 141 -----DPPPYE-PHC 150
DB 286 PCGPCSPCPCPYESPEC 301

RESULT 10
I37552
OX40 homologue - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-May-1998
C:Accession: I37552
R:Latza, U.; Durkop, H.; Schmittger, S.; Ringeling, J.; Eitelbach, F.; Hummel, M.; Fonat
Eur. J. Immunol. 24, 677-683, 1994
A:Title: The human OX40 homologue: cDNA structure, expression and chromosomal assignment c
A:Reference number: I37552; MUID:94170844
A:Accession: I37552
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-277 <RES>
A:Cross-references: EMBL:X75962; NID:9472957; PID:9472958

Query Match 13.3%; Score 116; DB 2; Length 277;
Best Local Similarity 28.6%; Pred. No. 0.0022;
Matches 30; Conservative 14; Mismatches 43; Indels 18; Gaps 6;

QY 50 CKQCGPHELSKEGFGVEDACVPCRPFRKEDWGFQKCPCADCALVNFQRAN-CS 108
DB 43 CHECPGNGMWSRC--SRSQNTVCPRCGGFYNDVYSSKPCPCPCWCNCRSGSERKOLCT 100

QY 109 HTSDAVCGDCLPG-----FYRTKTLVGFQDMECVPCGDPPEPYEP 148
DB 101 ATQDITVC-RCRAGTQPLDSYK-----PGVDCAPC--PPGHFSP 135

RESULT 11
GOVZML
T2 protein - myxoma virus (strain Lausanne)
C:Species: myxoma virus
C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 24-Oct-1997
```

```
C:Accession: A40566
R:Upton, C.; Macen, J.L.; Schreiber, M.; Mcfadden, G.
Virology 184, 370-382, 1991
A:Title: Myxoma virus expresses a secreted protein with homology to the tumor necrosi
A:Reference number: A40566; MUID:91335768
A:Accession: A40566
A:Molecule type: DNA
A:Residues: 1-326 <UPR>
A:Cross-references: GB:M95181; GB:M37976; NID:9332309; PID:9332310
C:Superfamily: myxoma virus T2 protein; NGF receptor repeat homology
C:Keywords: glycoprotein
F:64-105/Domain: NGF receptor repeat homology <NG2>
F:106-147/Domain: NGF receptor repeat homology <NG3>
F:166,181,205,238/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.2%; Score 115; DB 1; Length 326;
Best Local Similarity 29.7%; Pred. No. 0.003;
Matches 35; Conservative 9; Mismatches 60; Indels 14; Gaps 6;

QY 13 LEFAILFLHLAC-----KVSCEFGDCSRQGEFDRSGNCVLCQCGPHELSKEGFG 66
DB 1 MRLTLALVAVCYVGGAPYADGKRCGRNYED--GLC--CTSCPPGYSASRLC--G 54

QY 67 YGEDACVPCRPFRKEDWGF-QKCKPCADCALVNFQRANCSHTSDAVCGDCLPGFY 123
DB 55 PPSDITVCSFCKNETFTASTNHAPACVSCGRCTGHLSESQCDKTRDRVC--DCSAGNY 111

RESULT 12
A49053
CD27 antigen precursor - mouse
N:Alternate names: CD27L receptor; T cell activation antigen CD27
C:Species: Mus musculus (house mouse)
C:Date: 19-Dec-1993 #sequence_revision 22-Apr-1995 #text_change 11-Sep-1998
C:Accession: A49053
R:Gravesteln, L.A.; Blom, B.; Nolten, L.A.; de Vries, E.; van der Horst, G.; Ossendor
Eur. J. Immunol. 23, 943-950, 1993
A:Title: Cloning and expression of murine CD27: comparison with 4-1BB, another lympho
A:Reference number: A49053; MUID:93209296
A:Accession: A49053
A:Molecule type: mRNA
A:Residues: 1-250 <GRA>
A:Note: sequence extracted from NCBI backbone (NCBIN:128168, NCBIP:128169)
C:Superfamily: CD27 antigen; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; homodimer; receptor; surface antigen; T-cell;
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-250/Product: CD27 antigen #status predicted <MAT>
F:21-182/Domain: extracellular #status predicted <EXT>
F:27-63/Domain: NGF receptor repeat homology <NG1>
F:65-105/Domain: NGF receptor repeat homology <NG2>
F:121-179/Region: proline/serine/threonine-rich
F:183-202/Domain: transmembrane #status predicted <TMN>
F:203-250/Domain: intracellular #status predicted <INT>
F:95,162/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 13.1%; Score 114; DB 1; Length 250;
Best Local Similarity 28.1%; Pred. No. 0.003;
Matches 36; Conservative 15; Mismatches 41; Indels 36; Gaps 8;

QY 44 SGNVCVLCQCGPHELSKEGFGYED--ACQVPCRP-FRREKEDWGFQ-KKPCADCAL 98
DB 34 TGGGLCCNMCEGTFFVDC-----EQDRTAQCDCPCIGTSTSPPIYHPRHESCRHCN- 89

QY 99 VNRQRANCSHTSDAVCGDCLPGFYRTKTLVGFQDMECVPCG-----D 141
DB 89 -SGFLIRNCVTYANAE-SC-----SKNMQCRDQECTECDPINPALTRQPSFTSPQ 139

QY 142 PPPPYEPH 149
DB 140 PPTTLP 147
```

```

RESULT 13
GOHUN
nerv growth factor receptor precursor, low affinity - human
N:Alternate names: NGF receptor
C:Species: Homo sapiens (man)
C:Date: 31-Mar-1988 #sequence_revison 31-Mar-1988 #text_change 01-May-1998
C:Accession: A25218; A60204; S21689; I57638
R:Johnson, D.; Lanahan, A.; Buck, C.R.; Segal, A.; Morgan, C.; Mercer, E.; Bothwell, M.
Cell 47, 545-554, 1986
A:Title: Expression and structure of the human NGF receptor.
A:Reference number: A25218; M0ID: 87051725
A:Accession: A25218
A:Molecule type: mRNA
A:Residues: 1-427 <TOH>
A:Cross-references: GB:M14764; NID:g189204; PID:g189205
R:Marano, N.; Dietzschold, B.; Earley Jr., J.J.; Schattelman, G.; Thompson, S.; Grob, P.;
J. Neurochem. 48, 225-232, 1987
A:Title: Purification and amino terminal sequencing of human melanoma nerve growth factor.
A:Reference number: A60204; M0ID: 87085574
A:Accession: A60204
A:Molecule type: protein
A:Residues: 29-31, 'T', '33-42, 'TT', '45-46, 'TX', '50-51, 'XX', '54-56 <MAR>
A:Experimental source: melanoma cell line A875
A:Note: this sequence has been corrected by a note added in proof to follow the nucleoti
R:Visavajhala, P.; Leszyk, J.D.; Lin-Goecke, J.; Ross, A.H.
Arch. Biochem. Biophys. 294, 244-252, 1992
A:Title: Structural domains of the extracellular domain of human nerve growth factor rec
A:Reference number: S21689; M0ID: 92198017
A:Accession: S21689
A:Status: preliminary
A:Molecule type: protein
A:Residues: 183-208 <VIS>
R:Segal, A.; Patil, N.; Chao, M.
Mol. Cell. Biol. 8, 3160-3167, 1988
A:Title: A constitutive promoter directs expression of the nerve growth factor receptor
A:Reference number: I57638; M0ID: 89096903
A:Accession: I57638
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-22 <RNS>
A:Cross-references: GB:M21621; NID:g189206; PID:g189207
C:Comment: This receptor is found on sensory and sympathetic neurons, on neuroblastoma
C:Comment: The cysteine-rich region of the extracellular domain may form part or all of
C:Comment: This protein is thought to form a high-affinity receptor when it associates w
C:Comment: This receptor undergoes both N- and O-linked glycosylation.
C:Genetics:
A:Gene: GDB:NGFR
A:Cross-references: GDB:120234; OMIM:162010
A:Map position: 17q21-17q22
C:Superfamily: nerve growth factor receptor; NGF receptor repeat homology
C:Keywords: duplication; glycoprotein; heterodimer; monomer; phosphoprotein; receptor;
F:1-28/Domain: signal sequence #status predicted <SIG>
F:29-427/Product: nerve growth factor receptor #status experimental <MAT>
F:29-250/Domain: extracellular #status predicted <EXT>
F:32-67/Domain: NGF receptor repeat homology <NG1>
F:67-108/Domain: NGF receptor repeat homology <NG2>
F:109-147/Domain: NGF receptor repeat homology <NG3>
F:149-189/Domain: NGF receptor repeat homology <NG4>
F:197-248/Region: serine/threonine-rich
F:251-272/Domain: transmembrane #status predicted <TRM>
F:273-427/Domain: intracellular #status predicted <INT>
F:60/Binding site: carbohydrate (Asn) (covalent) #status predicted

```

```

Query Match 12.8%; Score 111.5; DB 1; Length 427;
Best Local Similarity 30.4%; Pred. No. 0.0074;
Matches 35; Conservative 10; Mismatches 49; Indels 21; Gaps 7;

```

```

OY 17 ILPLHL-----ACKVSCFEGDCSRQDEFDRSGNVCYLKQCGGMLSKEGCGYGEDA 71
DB 15 LLLLLGLVSLGAKACPTG-----LYTHSGEC--CRACNLGEGVADPCG---ANQT 62

```

```

OY 72 QCVPCRPFRFKED--WGFOCKRPACDCAALVNRFORANCSTSPAVCGDCLPGFYR 124
DB 63 VCEPCLDSVTSPDSVVSATPEKCKPCTECVGLQSNM-APCYEADDAVC-RCATGYTO 115

```

```

RESULT 14
B cell-associated surface molecule CD40 - mouse
A:Accession: A46515
C:Species: Mus musculus (house mouse)
C:Date: 18-Jun-1993 #sequence_revison 18-Nov-1994 #text_change 03-Mar-1995
C:Accession: A46515
R:Grimaldi, J.C.; Torres, R.; Kozak, C.A.; Chang, R.; Clark, E.A.; Howard, M.; Cockay
J. Immunol. 149, 3921-3926, 1992
A:Title: Genomic structure and chromosomal mapping of the murine CD40 gene.
A:Reference number: A46515; M0ID: 93094586
A:Accession: A46515
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: nucleic acid
A:Residues: 1-289 <GRI>
A:Experimental source: BALB/c, liver
A:Note: sequence extracted from NCBI backbone (NCBIP:120357)

```

```

Query Match 12.7%; Score 111; DB 2; Length 289;
Best Local Similarity 27.1%; Pred. No. 0.0062;
Matches 42; Conservative 9; Mismatches 58; Indels 46; Gaps 9;

```

```

OY 13 LFAILLFLHLACKVSCFEGDCSRQDEFDRSGNVCYLKQCGGMLSKEGCGYGEDAQ 72
DB 10 LMGCLTLAVHLG---OCVLT--CSDKQYLHD--GQC--CDLQPGSRLTSHC--TALEKTO 58
OY 73 CVPCCRPHRFKEDMG-----FQCKRPACDCAALVNRFORANCST--SPAVC----- 116
DB 59 CHPCDSGERSAQNREIRCHQHRHCEP-----NGLRKKEGTASDPTVCTCKEGQHC 111
OY 116 -----GDCLPGFYRKTLVGFQDMCEVPC 139
DB 112 TSKDCENCAQHTPCIPGFQVMEATETTTDTCVCHPC 146

```

```

RESULT 15
A46476
CD40 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 18-Jun-1993 #sequence_revison 18-Nov-1994 #text_change 16-Feb-1997
C:Accession: A46476
R:Torres, R.M.; Clark, E.A.
J. Immunol. 148, 620-626, 1992
A:Title: Differential increase of an alternatively polyadenylated mRNA species of mur
A:Reference number: A46476; M0ID: 92105763
A:Accession: A46476
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-305 <TOR>
A:Note: sequence extracted from NCBI backbone (NCBIN:75206, NCBIP:75207)
C:Keywords: transmembrane protein

```

```

Query Match 12.7%; Score 111; DB 2; Length 305;
Best Local Similarity 27.1%; Pred. No. 0.0064;
Matches 42; Conservative 9; Mismatches 58; Indels 46; Gaps 9;

```

```

OY 13 LFAILLFLHLACKVSCFEGDCSRQDEFDRSGNVCYLKQCGGMLSKEGCGYGEDAQ 72
DB 10 LMGCLTLAVHLG---OCVLT--CSDKQYLHD--GQC--CDLQPGSRLTSHC--TALEKTO 58
OY 73 CVPCCRPHRFKEDMG-----FQCKRPACDCAALVNRFORANCST--SPAVC----- 116
DB 59 CHPCDSGERSAQNREIRCHQHRHCEP-----NGLRKKEGTASDPTVCTCKEGQHC 111
OY 116 -----GDCLPGFYRKTLVGFQDMCEVPC 139
DB 112 TSKDCENCAQHTPCIPGFQVMEATETTTDTCVCHPC 146

```

Thu May 20 13:49:56 1999

us-09-212-270-4.rpr

Page 7

Search completed: May 13, 1999, 20:59:04
Job time: 2318 sec

This Page Blank (uspio)

This Page Blank (uspto)

| | |
|---|--|
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; |
| CC | EUTHERIA; RODENTIA. |
| RN | [1] |
| RP | SEQUENCE FROM N.A. |
| RC | STRAIN-BALB/C; |
| RX | MEDLINE: 94044750. |
| RA | CADHEAD D.M., BUHLMANN J.E., VAN DEN EERTWEGH A.J., |
| RA | CLAASSEN E., NOELLE R.J., FELT H.; |
| RL | J. IMMUNOL. 151:5261-5271(1993). |
| RN | [2] |
| RP | SEQUENCE FROM N.A. |
| RX | MEDLINE: 95255413. |
| RA | BIRKELAND M.L., COPELAND N.G., GILBERT D.J., JENKINS N.A., |
| RA | BARCLAY A.N.; |
| RL | EUR. J. IMMUNOL. 25:926-930(1995). |
| CC | - FUNCTION: RECEPTOR FOR THE OX40L/GP34 CYTOKINE. |
| CC | - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN. |
| DR | - SIMILARITY: CONTAINS A LA-NGRF/TNFR-TYPE CYSTEINE-RICH REGION, |
| DR | EMBL: Z21674; G312828; - |
| DR | EMBL: X85214; G732819; - |
| DR | MGI: MGI:104512; TXGDI. |
| DR | PROSITE: PS00652; TNFR_NGFR.1; 3. |
| DR | PROSITE: PS00650; TNFR_NGFR.2; 2. |
| KM | RECEPTOR; T-CELL; ANTIGEN; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; |
| KM | SIGNAL. |
| FT | FT SIGNAL. 1 19 POTENTIAL. |
| FT | FT CHAIN 20 272 OX40L RECEPTOR. |
| FT | FT DOMAIN 20 211 EXTRACELLULAR (POTENTIAL). |
| FT | FT TRANSMEM 212 236 POTENTIAL. |
| FT | FT DOMAIN 237 272 CYTOPLASMIC (POTENTIAL). |
| FT | FT DOMAIN 26 165 4 x TNFR-CYS. |
| FT | FT REPEAT 26 61 TNFR-CYS 1. |
| FT | FT REPEAT 62 103 TNFR-CYS 2. |
| FT | FT REPEAT 104 124 TNFR-CYS 3 (INCOMPLETE). |
| FT | FT REPEAT 125 165 TNFR-CYS 4. |
| FT | FT CARBOHYD 144 144 POTENTIAL. |
| FT | FT CONFLICT 15 15 A->G (IN REF. 2). |
| SO | SEQUENCE 272 AA; 30153 MW; 67DIB978 CRC32; |
| <hr/> | |
| Query Match 15.5%; Score 135.5; DB 1; Length 272; | |
| Best Local Similarity 29.4%; Pred. No. 1.8e-05; | |
| Matches 35; Conservative 14; Mismatches 43; Indels 27; Gaps: | |
| Oy | 46 NCV-----LCQCGFMELSRCEGFYGGEAQCVCPRPHRKFKDMGFCCKPCADC 966 Db 26 NCNVHTYPSGHKKCRCCRCRCPCHGMVSR--DHTRDLTCHPCETGYENAVNYDKCKQTCC 833 |
| Oy | 97 ALVNRFQ-RANGSHTESDAYCSCGLTGFRKTKLVFOD-----MECVPGCDPPPYEP 148 Db 84 NHNSGSLSKONCTPTPDIVC-RCRGTQDP-----QDSGKGIGADVCPV--PPGHFSF 1333 |
| <hr/> | |
| RESULT 3 | |
| ID LMB1_HUMAN | STANDARD; PRJ: 1786 AA. |
| AC P07942: | |
| DT 01-AUG-1988 (REL. 08. CREATED) | |
| DT 01-AUG-1988 (REL. 08. LAST SEQUENCE UPDATE) | |
| DT 15-JUL-1998 (REL. 36. LAST ANNOTATION UPDATE) | |
| DE LAMININ BETA-1 CHAIN PRECURSOR (LAMININ B1 CHAIN). | |
| GN LAMB1. | |
| OS HOMO SAPIENS (HUMAN) | |
| OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA; | |
| OC EUTHERIA; PRIMATES. | |
| RN [1] | |
| RP SEQUENCE FROM N.A. | |
| RX MEDLINE: 90368768. | |
| RA VIOLTEENAHO R., CHOW L.T., TRYGGVASON K.; | |
| RL J. BIOL. CHEM. 265:15611-15616(1990). | |
| [2] | |
| RP SEQUENCE FROM N.A. | |
| RX MEDLINE: 87280097. | |

RA PIKKAIKAIEN T., EDDY R., FUKUSHIMA Y., BYERS M., SHONS T.,
RA PILAIANIANIEN T., SARASTE M., TRYGGVASON K.,
RN J. BIOL. CHEM. 262:10454-10462(1987).
[3]
RN SEQUENCE OF 1276-1709 FROM N.A.
RP MEDLINE; 88021029.
RA JAVE N., MOOI W.S., RICCA G.A., MOOD R., CHIU I.M., O'BRIEN S.J.,
RA DROHAN W.N.,
RL AM. J. HUM. GENET. 41:605-615(1987).
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC -1- SUBUNIT: THE BETA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ),
CC LAMININ-2 (MEROSIN), AND LAMININ-6 (K-LAMININ).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR
CC COMPONENT).
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.
DR EMBL; M61951; G186913; -
DR EMBL; M58147; G186913; JOINED.
DR EMBL; M61917; G186913; JOINED.
DR EMBL; M61918; G186913; JOINED.
DR EMBL; M61921; G186913; JOINED.
DR EMBL; M61922; G186913; JOINED.
DR EMBL; M61923; G186913; JOINED.
DR EMBL; M61924; G186913; JOINED.
DR EMBL; M61925; G186913; JOINED.
DR EMBL; M61926; G186913; JOINED.
DR EMBL; M61927; G186913; JOINED.
DR EMBL; M61928; G186913; JOINED.
DR EMBL; M61929; G186913; JOINED.
DR EMBL; M61930; G186913; JOINED.
DR EMBL; M61931; G186913; JOINED.
DR EMBL; M61932; G186913; JOINED.
DR EMBL; M61933; G186913; JOINED.
DR EMBL; M61934; G186913; JOINED.
DR EMBL; M61935; G186913; JOINED.
DR EMBL; M61936; G186913; JOINED.
DR EMBL; M61938; G186913; JOINED.
DR EMBL; M61939; G186913; JOINED.
DR EMBL; M61940; G186913; JOINED.
DR EMBL; M61941; G186913; JOINED.
DR EMBL; M61942; G186913; JOINED.
DR EMBL; M61943; G186913; JOINED.
DR EMBL; M61944; G186913; JOINED.
DR EMBL; M61945; G186913; JOINED.
DR EMBL; M61946; G186913; JOINED.
DR EMBL; M61947; G186913; JOINED.
DR EMBL; M61948; G186913; JOINED.
DR EMBL; M61949; G186913; JOINED.
DR EMBL; M61950; G186913; JOINED.
DR EMBL; M55370; G186876; -
DR EMBL; M55378; G186876; JOINED.
DR EMBL; M55365; G186876; JOINED.
DR EMBL; M55371; G186876; JOINED.
DR EMBL; M55372; G186876; JOINED.
DR EMBL; M55373; G186876; JOINED.
DR EMBL; M55374; G186876; JOINED.
DR EMBL; M55375; G186876; JOINED.
DR EMBL; M55376; G186876; JOINED.
DR EMBL; M55344; G186876; JOINED.
DR EMBL; M55345; G186876; JOINED.
DR EMBL; M55346; G186876; JOINED.

DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE LYMPHOTOXIN-BETA RECEPTOR PRECURSOR (TUMOR NECROSIS FACTOR RECEPTOR
 DE 2 RELATED PROTEIN) (TUMOR NECROSIS FACTOR C RECEPTOR).
 GN LTR OR TNECR.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=LIVER.
 RX MEDLINE: 93252381.
 RA BANS M., CHAFFNET M., CASSIMAN J.J., DEN BERGHE H., MARIEN P.,
 RL GENOMICS 16:214-218(1993).
 RN [2]
 RP FUNCTION.
 RX MEDLINE: 94225209.
 RA CROME P.D., VAN ARSDALE T.L., WALTER B.N., WARE C.F., HESSION C.,
 RA EHRENFELS B., BROWNING J.L., DIN W.S., GOODWIN R.G., SMITH C.A.;
 RL SCIENCE 264:707-710(1994).
 CC -1- FUNCTION: RECEPTOR FOR THE LYMPHOTOXIN-BETA. POSSIBLE FUNCTION IN
 CC IMMUNE DEVELOPMENT.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: L04270; G339762; -.
 DR HSSP: P19999; ICLG.
 DR MIM: 600979; -.
 DR PROSITE: PS00653; TNFR_NGFR_1; 2.
 DR PROSITE: PS50050; TNFR_NGFR_2; 3.
 KW RECEPTOR; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT; SIGNAL.
 FT SIGNAL 1 30
 FT CHAIN 1 435
 FT DOMAIN 31 227
 FT TRANSEM 228 248
 FT DOMAIN 249 435
 FT DOMAIN 42 211
 FT REPEAT 42 81
 FT REPEAT 82 124
 FT REPEAT 125 168
 FT REPEAT 169 211
 FT DISULFID 43 58
 FT DISULFID 59 72
 FT DISULFID 72 80
 FT DISULFID 83 98
 FT DISULFID 101 116
 FT DISULFID 104 124
 FT DISULFID 126 132
 FT DISULFID 139 148
 FT DISULFID 142 167
 FT DISULFID 170 185
 FT CARBOHYD 40 40
 FT CARBOHYD 177 177
 SQ SEQUENCE 435 AA; 46709 MW; 203882DD CRC32;

WSL1_HUMAN
 ID WSL1_HUMAN STANDARD: PRT; 417 AA.
 AC 093038; 093036; 093037; 092983; P78515; Q99831; Q99722; P78507;
 DT 01-NOV-1987 (REL. 35, CREATED)
 DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)
 DT 01-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE WSL-1 PROTEIN PRECURSOR (APOPTOSIS-MEDIATING RECEPTOR DR3) (APOPTOSIS-
 DE MEDIATING RECEPTOR TRAMP) (DEATH DOMAIN RECEPTOR 3) (WSL PROTEIN)
 DE (APOPTOSIS INDUCING RECEPTOR AIR) (APO-3) (LYMPHOCYTE ASSOCIATED
 DE RECEPTOR OF DEATH) (LARD)
 GN WSL1 OR WSL OR DR3 OR DR3.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND MUTAGENESIS.
 RC TISSUE=LYMPHOID;
 RX MEDLINE: 97088617.
 RA KITSON J., RAVEN T., JIANG Y.-P., GOEDDEL D.V., GILES K.M., PUN K.-T.,
 RA GRINHAM C.J., BROWN R., FARROW S.N.;
 RL NATURE 384:372-375(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=UMBILICAL VEIN ENDOTHELIAL CELLS;
 RX MEDLINE: 97081063.
 RA CHINNAIVAN A.M., O'ROURKE K., YU G.-L., LYONS R.H., GARG M.,
 RA DUAN D.R., XING L., GENTZ R., NI J., DIXIT V.M.;
 RL SCIENCE 274:990-992(1996).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA DEGLI-ESPOSTI M.A., DIN W.S., COSMAN D., SMITH C.A., GOODWIN R.G.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=HEART;
 RA MARSTERS S.A., SHERIDAN J.P., DONAHUE C.J., PITTI R.M., GRAY C.L.,
 RA GODDARD A.D., BAUER K.D., ASHKENAZI A.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA SKEATON G., XU X.N., OLSEN A., COMPER A., TAN R., MCMICHAEL A.,
 RA BELL J.;
 RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [6]
 RP SEQUENCE OF 4-417 FROM N.A.
 RC TISSUE-BRAIN, AND FETAL LUNG;
 RX MEDLINE: 97205335.
 RA BODMER J.-L., BURNS K., SCHNEIDER P., HOFMANN K., STEINER V.,
 RA THOME M., BORNARD T., HAHNE M., SCHROETER M., WILSON A., FRENCH L.E.,
 RA BROWNING J.L., MACDONALD H.R., TSCHOPP J.;
 RL IMMUNITY 6:79-88(1997).
 RN [7]
 RP SEQUENCE OF 7-417 FROM N.A.
 RC TISSUE-BRAIN;
 RA CHAUDHARY P.M., HOOD L.E.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: INDUCES APOPTOSIS AND ACTIVATES NUCLEAR FACTOR KAPPA-B
 CC (NF-KAPPA-B). DIRECTLY INTERACTS WITH TRADD ADAPTOR MOLECULE. MAY
 CC PLAY A ROLE IN REGULATING LYMPHOCYTE HOMEOSTASIS.
 CC -1- SUBUNIT: HOMODIMER. INTERACTS STRONGLY VIA THE DEATH DOMAINS WITH
 CC THE TNFR1-ASSOCIATED MOLECULE TRADD AND THE TNFR1 RECEPTOR TO
 CC ACTIVATE AT LEAST TWO DISTINCT SIGNALING CASCADES, APOPTOSIS AND
 CC NF-KAPPA B SIGNALING.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN (POTENTIAL).
 CC -1- TISSUE SPECIFICITY: ABUNDANTLY EXPRESSED IN THYMOCYTES AND
 CC LYMPHOCYTES. DETECTED IN LYMPHOCYTE-RICH TISSUES SUCH AS THYMUS,
 CC COLON, INTESTINE, AND SPLEEN. ALSO FOUND IN THE PROSTATE.
 CC -1- PTM: GLYCOSYLATED (PROBABLE).
 CC -1- ALTERNATIVE PRODUCTS: THREE ISOFORMS WSL-1/LARD-1A (SHOWN HERE),
 CC WSL-S1/LARD-3, AND WSL-S2 ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 DR EMBL: Y09392; E281462; -.
 DR EMBL: Y09392; E281460; -.

```
DR EMBL: Y09392; E281461; -
DR EMBL: U72763; G1669512; -
DR EMBL: U83599; G1800297; -
DR EMBL: U83600; G1800299; -
DR EMBL: U78029; G1778764; -
DR EMBL: U74611; G1763293; -
DR EMBL: U94501; G2071949; -
DR EMBL: U94504; G2071955; -
DR EMBL: U73580; G1695925; -
DR EMBL: U73581; G1695927; -
DR EMBL: U83597; G1800293; -
DR PROSITE: PS00652; TNER_NGFR.1; 2.
DR PROSITE: PS50050; TNER_NGFR.2; 1.
DR PROSITE: PS50017; DEATH_DOMAIN.1.
KW RECEPTOR; APOPTOSIS; ALTERNATIVE SPLICING; TRANSMEMBRANE; SIGNAL;
KW REPEAT.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 417 WSL-1 PROTEIN.
FT DOMAIN 25 199 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 200 220 POTENTIAL.
FT DOMAIN 223 417 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 34 192 4 X TNER-CYS.
FT REPEAT 34 71 TNER-CYS 1.
FT REPEAT 72 115 TNER-CYS 2.
FT REPEAT 116 163 TNER-CYS 3.
FT REPEAT 164 192 TNER-CYS 4.
FT DOMAIN 332 413 DEATH DOMAIN.
FT CARBOHYD 67 67 POTENTIAL.
FT CARBOHYD 106 106 POTENTIAL.
FT VARSPLIC 182 218 STIGSPERCAAVCGRMFWOVLLAGLVPLILGA ->
VIGBGRPCGPRPANHPLDHPPLASQAPGICR (IN
FT VARSPLIC 219 417 WSL-S1/LARD-3).
FT VARSPLIC 200 253 MISSING (IN WSL-S1/LARD-3).
FT VARSPLIC 200 253 MEMVOVLLAGLVPLILGATLTYTRHCGMPKPLVTADGAG
MEALTPPATHTLS -> SRMCAGNARGGTGMDRGAGEGCG
NHPTPTSCFGCGSGRCSMLALMSPECLAPP (IN WSL-
FT VARSPLIC 254 417 MISSING (IN WSL-S2).
FT MUTAGEN 354 354 L->A: SUPPRESSES HOMODIMERIZATION, TNER1
INTERACTION, AND APOPTOSIS INDUCTION.
FT MUTAGEN 356 356 L->A: SUPPRESSES HOMODIMERIZATION, AND
TNER1 INTERACTION.
FT MUTAGEN 373 373 D->A: SUPPRESSES HOMODIMERIZATION, AND
TNER1 INTERACTION.
FT CONFLICT 4 6 RPR -> AAA (IN REF. 6).
FT CONFLICT 60 60 P -> H (IN REF. 7).
FT CONFLICT 167 167 P -> L (IN REF. 6 AND 7).
FT CONFLICT 312 312 A -> R (IN REF. 1).
FT CONFLICT 370 370 R -> L (IN REF. 1).
FT CONFLICT 381 381 R -> H (IN REF. 1).
FT SEQUENCE 417 AA: 45385 MW: 3AEEFOA5 CRC32;

Query Match 14.1%; Score 123.5; DB 1: Length 417;
Best Local Similarity 31.9%; Pred. No. 0.00029;
Matches 37; Conservative 12; Mismatches 34; Indels 33; Gaps 8;
```

```
DE LAMININ BETA-2 CHAIN PRECURSOR.
GN LAMB2 OR S-LAM OR LAMS.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-129/J;
RX MEDLINE: 96278760.
RA DURKIN M.E., GAUTAM M., LOEHEL S., SANES J.R., MERLIE J.P.,
RA ALBRECHTSEN R., WEMER U.M.;
RL J. BIOL. CHEM. 271:13407-13416(1996).
RN [2]
RP SEQUENCE OF 348-428 FROM N.A.
RC TISSUE-LUNG;
RX MEDLINE: 94319092.
RA ABERDAM D., GALLIANO M.F., MATTEI M.-G., ORTONNE J.P., MENEGUZZI G.;
RL MAMM. GENOME 5:393-394(1994).
RN [3]
RP FUNCTION.
RC STRAIN-129/J;
RX MEDLINE: 95191650.
RA NOAKES P.G., GAUTAM M., MUDD J., SANES J.R., MERLIE J.P.;
RL NATURE 374:258-262(1995).
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ
IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF
CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING
WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.
CC -1- FUNCTION: LAMININ-3 (S-LAMININ) REGULATES THE FORMATION OF MOTOR
NERVE TERMINALS.
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE
DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND
TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE
COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.
CC -1- SUBUNIT: THE BETA-2 CHAIN IS A SUBUNIT OF LAMININ-3 (S-LAMININ),
LAMININ-4 (S-MEROSIN), AND LAMININ-7 (KS-LAMININ).
CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR; FOUND IN THE BASEMENT
MEMBRANES (MAJOR COMPONENT).
CC -1- TISSUE SPECIFICITY: NEUROMUSCULAR SYNAPSE AND KIDNEY GLOMERULUS.
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT
WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.
CC -1- DOMAIN: DOMAINS VI AND IV ARE GLOBULAR.
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).
CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.
DR EMBL: U43541; G1244720; -.
DR EMBL: U42624; G1244720; JOINED.
DR EMBL: X75928; G511042; -.
DR MGD: MGI:99916; LAMB2.
DR PROSITE: PS00022; EGF_1; 10.
DR PROSITE: PS01186; EGF_2; 2.
DR PROSITE: PS01248; LAMININ TYPE EGF; 12.
KW GLYCOPROTEIN; BASEMENT MEMBRANE; EXTRACELLULAR MATRIX; COILED COIL;
KW LAMININ EGF-LIKE DOMAIN; CELL ADHESION; REPEAT; SIGNAL.
FT SIGNAL 1 35 POTENTIAL.
FT CHAIN 36 1799 LAMININ BETA-2 CHAIN.
FT DOMAIN 36 283 LAMININ N-TERMINAL (DOMAIN VI).
FT DOMAIN 284 555 4.5 X LAMININ EGF-LIKE REPEATS (DOMAIN
V).
FT DOMAIN 286 349 LAMININ EGF-LIKE 1.
FT DOMAIN 350 412 LAMININ EGF-LIKE 2.
FT DOMAIN 413 472 LAMININ EGF-LIKE 3.
FT DOMAIN 473 524 LAMININ EGF-LIKE 4.
FT DOMAIN 525 555 LAMININ EGF-LIKE 5 (INCOMPLETE).
FT DOMAIN 556 782 LAMININ DOMAIN IV.
FT DOMAIN 783 1191 8 X LAMININ EGF-LIKE REPEATS (DOMAIN
III).
FT DOMAIN 784 831 LAMININ EGF-LIKE 6.
FT DOMAIN 832 877 LAMININ EGF-LIKE 7.
FT DOMAIN 878 927 LAMININ EGF-LIKE 8.
FT DOMAIN 928 986 LAMININ EGF-LIKE 9.
FT DOMAIN 987 1038 LAMININ EGF-LIKE 10.
FT DOMAIN 1039 1095 LAMININ EGF-LIKE 11.
```


FT DOMAIN 541 772 LAMININ DOMAIN IV.
 FT DOMAIN 773 1178 8 X LAMININ EGF-LIKE REPEATS (DOMAIN
 FT DOMAIN 773 820 III).
 FT DOMAIN 821 866 LAMININ EGF-LIKE 6.
 FT DOMAIN 867 916 LAMININ EGF-LIKE 7.
 FT DOMAIN 917 975 LAMININ EGF-LIKE 8.
 FT DOMAIN 976 1027 LAMININ EGF-LIKE 9.
 FT DOMAIN 1028 1083 LAMININ EGF-LIKE 10.
 FT DOMAIN 1084 1131 LAMININ EGF-LIKE 11.
 FT DOMAIN 1132 1178 LAMININ EGF-LIKE 12.
 FT DOMAIN 1179 1397 LAMININ EGF-LIKE 13.
 FT DOMAIN 1398 1430 DOMAIN II.
 FT DOMAIN 1431 1786 DOMAIN ALPHA.
 FT DOMAIN 1216 1315 COILED COIL (POTENTIAL).
 FT DOMAIN 1368 1388 COILED COIL (POTENTIAL).
 FT DOMAIN 1448 1778 COILED COIL (POTENTIAL).
 FT DISULFID 271 280 BY SIMILARITY.
 FT DISULFID 273 298 BY SIMILARITY.
 FT DISULFID 300 309 BY SIMILARITY.
 FT DISULFID 312 332 BY SIMILARITY.
 FT DISULFID 335 344 BY SIMILARITY.
 FT DISULFID 337 362 BY SIMILARITY.
 FT DISULFID 365 374 BY SIMILARITY.
 FT DISULFID 377 395 BY SIMILARITY.
 FT DISULFID 398 411 BY SIMILARITY.
 FT DISULFID 400 426 BY SIMILARITY.
 FT DISULFID 428 437 BY SIMILARITY.
 FT DISULFID 440 455 BY SIMILARITY.
 FT DISULFID 458 472 BY SIMILARITY.
 FT DISULFID 479 490 BY SIMILARITY.
 FT DISULFID 481 490 BY SIMILARITY.
 FT DISULFID 493 507 BY SIMILARITY.
 FT DISULFID 773 785 BY SIMILARITY.
 FT DISULFID 775 792 BY SIMILARITY.
 FT DISULFID 794 803 BY SIMILARITY.
 FT DISULFID 806 818 BY SIMILARITY.
 FT DISULFID 821 833 BY SIMILARITY.
 FT DISULFID 823 840 BY SIMILARITY.
 FT DISULFID 842 851 BY SIMILARITY.
 FT DISULFID 854 864 BY SIMILARITY.
 FT DISULFID 867 876 BY SIMILARITY.
 FT DISULFID 869 883 BY SIMILARITY.
 FT DISULFID 886 895 BY SIMILARITY.
 FT DISULFID 898 914 BY SIMILARITY.
 FT DISULFID 917 933 BY SIMILARITY.
 FT DISULFID 919 944 BY SIMILARITY.
 FT DISULFID 946 955 BY SIMILARITY.
 FT DISULFID 958 973 BY SIMILARITY.
 FT DISULFID 976 990 BY SIMILARITY.
 FT DISULFID 978 997 BY SIMILARITY.
 FT DISULFID 1000 1009 BY SIMILARITY.
 FT DISULFID 1012 1025 BY SIMILARITY.
 FT DISULFID 1084 1096 BY SIMILARITY.
 FT DISULFID 1086 1103 BY SIMILARITY.
 FT DISULFID 1105 1114 BY SIMILARITY.
 FT DISULFID 1117 1129 BY SIMILARITY.
 FT DISULFID 1132 1144 BY SIMILARITY.
 FT DISULFID 1134 1151 BY SIMILARITY.
 FT DISULFID 1153 1162 BY SIMILARITY.
 FT DISULFID 1165 1176 BY SIMILARITY.
 FT DISULFID 1179 1179 INTERCHAIN (PROBABLE).
 FT DISULFID 1182 1182 INTERCHAIN (PROBABLE).
 FT DISULFID 1185 1785 INTERCHAIN (PROBABLE).
 FT DISULFID 120 120 POTENTIAL.
 FT CARBOHYD 356 356 POTENTIAL.
 FT CARBOHYD 519 519 POTENTIAL.
 FT CARBOHYD 677 677 POTENTIAL.
 FT CARBOHYD 1041 1041 POTENTIAL.
 FT CARBOHYD 1195 1195 POTENTIAL.
 FT CARBOHYD 1279 1279 POTENTIAL.
 FT CARBOHYD 1336 1336 POTENTIAL.
 FT CARBOHYD 1343 1343 POTENTIAL.

FT CARBOHYD 1487 1487 POTENTIAL.
 FT CARBOHYD 1533 1533 POTENTIAL.
 FT CARBOHYD 1542 1542 POTENTIAL.
 FT CARBOHYD 1643 1643 POTENTIAL.
 FT CONFLICT 1531 1534 POTENTIAL.
 FT CONFLICT 1749 1749 D -> N (IN REF. 2).
 SQ SEQUENCE 1786 AA; 196904 MW; 56CA7685 CRC32;
 Query Match 13.7%; Score 120; DB 1; Length 1786;
 Best Local Similarity 24.5%; Pred. No. 0.0018;
 Matches 45; Conservative 18; Mismatches 57; Indels 64; Gaps 11;
 Oy 10 RTVEAALFLFLH---LACKVS-----CET--GDSCROEFGKRSNCVLCXQCGP 56
 Db 756 RNIIF-SISALIHQTGLACEDPDGSSVCDPNNGCQCRRNVVGRY-----CNRCAPG 809
 Oy 57 MELSKECGFYGEDACVPCRPNR-----FKEADGCF 87
 Db 810 T-----FGGPNNG-CKPCDCHLGSASAFCDATGCHCFQGIYARQCDRLPGYWG 861
 Oy 88 QKCPK-----ADCAVNRFPORANCSTSDAVCGDCLPGFYRKTLYGFQDMECVPCGD 141
 Db 862 PSCPCQCNCHALDDCDYTGECCLSCODYTGTGHNCEKCLAGY-GDPIIGSGD-HCRPCPC 919
 Oy 142 PPPP 145
 Db 920 PDGP 923
 RESULT 8
 CD40_HUMAN STANDARD; PRT; 277 AA.
 ID CD40_HUMAN P25942;
 AC P25942;
 DT 01-MAY-1992 (REL. 22, CREATED)
 DT 01-MAY-1992 (REL. 22, LAST SEQUENCE UPDATE)
 DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
 DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40).
 GN CD40.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 89356608.
 RA STAMENKOVIC I., CLARK E.A., SEED B.;
 RL EMBO J. 8:1403-1410(1989).
 RN [2]
 RP 3D-STRUCTURE MODELLING OF 24-144.
 RX MEDLINE; 97189482.
 RA BAJORATH J., ARUFFO A.;
 RL PROTEINS 27:59-70(1997).
 RN [3]
 RP 3D-STRUCTURE MODELLING OF 26-186 IN COMPLEX WITH CD40L.
 RX MEDLINE; 98266353.
 RA SINGH J., GARBER E., VAN VLIJMEN H., KARPSUSAS M., HSU Y.-M.,
 RA ZHENG Z., NAISWORTH J.H., THOMAS D.;
 RL PROTEIN SCI. 7:1124-1135(1998).
 CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
 CC -1- TISSUE SPECIFICITY: B-CELLS AND IN PRIMARY CARCINOMAS.
 CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
 CC EMBL; X60592; E1175755; -.
 DR PIR; S04460; S04460.
 DR PDB; 1CDF; 01-APR-97.
 DR MIM; 109535; -.
 DR PROSITE; PS00652; TNFR_NGFR_1; 1.
 DR PROSITE; PS50050; TNFR_NGFR_2; 4.
 KW RECEPTOR; B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL;
 FT 3D-STRUCTURE.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 277 CD40L RECEPTOR.
 FT DOMAIN 20 193 EXTRACELLULAR (POTENTIAL).


```

OY 67 YGEDACQVPCRPHEKEDMGF-QKCKPCADCALVNRFORANCSTDAVCGDCLPFGFY 123
DB 55 PGGDVCPCNKETFTASTNHPACVSGRGCTGHLSESQCDKTRDYC-DCSAGNY 111

RESULT 13
CD27_MOUSE STANDARD: PRT: 250 AA.
ID CD27_MOUSE
AC P41272:
DT 01-FEB-1995 (REL. 31, CREATED)
DT 01-FEB-1995 (REL. 31, LAST SEQUENCE UPDATE)
DE 01-FEB-1995 (REL. 31, LAST ANNOTATION UPDATE)
DE CD27L RECEPTOR PRECURSOR (T-CELL ACTIVATION ANTIGEN CD27).
GN CD27.
OS MUS MUSCULUS (MUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-B6/CBA; TISSUE-THYMUS;
RX MEDLINE; 93209296.
RA GRAVESTEIN L.A., BLOM B., NOLTEN L.A., DE VRIES E., VAN DER HORST G.,
RA OSSENDOOP F., BORST J., LOENEN W.A.;
RL EUR. J. IMMUNOL. 23:943-950(1993).
CC -1- FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD27L. MAY PLAY
CC -1- A ROLE IN SURVIVAL OF ACTIVATED T-CELLS.
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- TISSUE SPECIFICITY: IN THYMUS AND SPLEEN, BUT NOT IN NON-LYMPHOID
CC TISSUES.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: L24495; -: NOT_ANNOTATED_CDS.
DR PIR: A49053; A49053.
DR MGD; MGI:88326; CD27.
DR PROSITE; PS00652; TNFR_NGFR_1; 1.
DR PROSITE; PS50050; TNFR_NGFR_2; 1.
KW T-CELL; GLYCOPROTEIN; TRANSMEMBRANE; SIGNAL; RECEPTOR; REPEAT.
FT SIGNAL 1 20 POTENTIAL.
FT CHAIN 21 250 CD27L RECEPTOR.
FT DOMAIN 21 182 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 183 203 POTENTIAL.
FT DOMAIN 204 250 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 26 141 3 X TNFR-CYS.
FT REPEAT 26 63 TNFR-CYS 1.
FT REPEAT 64 104 TNFR-CYS 2.
FT REPEAT 105 141 TNFR-CYS 3.
FT CARBOHYD 95 95 POTENTIAL.
FT CARBOHYD 162 162 POTENTIAL.
SQ SEQUENCE 250 AA; 28164 MW; AFEF74ED CRC32;

Query Match 13.1%; Score 114; DB 1; Length 250;
Best Local Similarity 28.1%; Pred. No. 0.0013;
Matches 36; Conservative 15; Mismatches 41; Indels 36; Gaps 8;
OY 44 SGNVCYLKQCGGMLSKGCGGYGSD--ACQVPCRP-HRKEMWGRQ-KKPPADCAL 98
DB 34 TGGGACRCRCGEGTFVVDK-----BDDRTAACQDCPCIPGTSTSPDYHTRPHCSCRHCN- 89
OY 99 VNRFORANCSTDAVCGDCLPGRYRTKLVGFQDMCEVPCG-----D 141
DB 89 -SGFLIRNCTVYANBC-SC-----SKNMCCRQCECTECPLPALTRQPSSTPSQ 139
OY 142 PPPPYEPH 149
DB 140 PPTPLPH 147

RESULT 14
NGFR_HUMAN STANDARD: PRT: 427 AA.
ID NGFR_HUMAN
AC P08138;
DT 01-AUG-1988 (REL. 08, CREATED)

```

```

DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)
DT 15-JUN-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE LOW-AFFINITY NERVE GROWTH FACTOR RECEPTOR PRECURSOR (NGF RECEPTOR)
DE (GP80-LINGER) (P75 ICD).
GN NGFR.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE; 87051725.
RA JOHNSON D., LANAHAN A., BUCK C.R., SEHGAL A., MORGAN C., MERCER E.,
RA BOTHWELL M., CHAO M.;
RL CELL 47:545-554(1986).
CC -1- FUNCTION: LOW AFFINITY RECEPTOR WHICH CAN BIND TO NGF, BDNF,
CC NT-3, AND NT-4.
CC -1- SUBUNIT: NGF RECEPTOR CAN FORM A HOMODIMER THROUGH DISULFIDE
CC BOND FORMATION.
CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC -1- PTM: N- AND O-GLYCOSYLATED AND IS PHOSPHORYLATED ON SERINE.
CC -1- SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL; M14764; G189205; -.
DR PIR; A25218; GQHUN.
DR HSSP; P19438; ITNR.
DR MIM; 162010; -.
DR PROSITE; PS00652; TNFR_NGFR_1; 3.
DR PROSITE; PS50050; TNFR_NGFR_2; 4.
DR PROSITE; PS50017; DEATH_DOMAIN; 1.
KW RECEPTOR; NEUROGENESIS; TRANSMEMBRANE; GLYCOPROTEIN; REPEAT;
KW PHOSPHORYLATION; SIGNAL.
FT SIGNAL 1 28
FT CHAIN 29 427 NGF RECEPTOR.
FT DOMAIN 29 250 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 251 272 POTENTIAL.
FT DOMAIN 273 427 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 31 189 4 X TNFR-CYS.
FT REPEAT 31 65 TNFR-CYS 1.
FT REPEAT 66 107 TNFR-CYS 2.
FT REPEAT 108 147 TNFR-CYS 3.
FT REPEAT 148 189 TNFR-CYS 4.
FT DOMAIN 344 421 DEATH DOMAIN.
FT DISULFID 32 43 BY SIMILARITY.
FT DISULFID 44 57 BY SIMILARITY.
FT DISULFID 47 64 BY SIMILARITY.
FT DISULFID 67 83 BY SIMILARITY.
FT DISULFID 86 99 BY SIMILARITY.
FT DISULFID 89 107 BY SIMILARITY.
FT DISULFID 109 122 BY SIMILARITY.
FT DISULFID 125 138 BY SIMILARITY.
FT DISULFID 128 146 BY SIMILARITY.
FT DISULFID 149 164 BY SIMILARITY.
FT DISULFID 167 180 BY SIMILARITY.
FT DISULFID 170 188 BY SIMILARITY.
FT DOMAIN 197 248 SER/THR-RICH.
FT CARBOHYD 60 60 POTENTIAL.
SQ SEQUENCE 427 AA; 45183 MW; BE2924BD CRC32;

Query Match 12.8%; Score 111.5; DB 1; Length 427;
Best Local Similarity 30.4%; Pred. No. 0.0033;
Matches 35; Conservative 10; Mismatches 49; Indels 21; Gaps 7;
OY 17 ILFLHL-----ACKVSCETGDCSRQDFKRSNGVCYLKQCGGMLSKGCGGYGSDA 71
DB 15 LLLLLLVSLGAKKACPTG-----LYTHSGEC-CKKCNINSGEVAQPCG--ANQT 62
OY 72 QCVPCRPHEKED--WGRQCKPCADCALVNRFORANCSTDAVCGDCLPFGFYR 124
DB 63 VCEPCLDSVTFSDVVASATPECKPCTCEVGLOSMS-APCEVADDAVC-RCAYGYQ 115

RESULT 15
CD40_MOUSE

```

Search completed: May 19, 1999, 14:43:39
Job time: 46 sec

ID CD40_MOUSE STANDARD; PRT; 289 AA.
AC P27512;
DT 01-AUG-1992 (REL. 23, CREATED)
DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE CD40L RECEPTOR PRECURSOR (B-CELL SURFACE ANTIGEN CD40) (BP50) (CDW40).
GN CD40.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92105763.
RA TORRES R.M., CLARK E.A.;
RL J. IMMUNOL. 148:620-626(1992).
RN [2]
RP REVISIONS.
RC STRAIN-BALB/C;
RA TORRES R.M.;
RL SUBMITTED (SEP-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C; TISSUE=LIVER;
RX MEDLINE: 93094586.
RA GRIMALDI J.C., TORRES R., KOZAK C.A., CHANG R., CLARK E.A.,
RA HOWARD M., COCKAYNE D.A.;
RL J. IMMUNOL. 149:3921-3926(1992).
CC - FUNCTION: RECEPTOR FOR A CYTOKINE LIGAND KNOWN AS CD40L.
CC - SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.
CC - SIMILARITY: CONTAINS A LA-NGFR/TNFR-TYPE CYSTEINE-RICH REGION.
DR EMBL: M83312; G1553059; .
DR EMBL: M94126; G192526; .
DR EMBL: M94129; G192526; JOINED.
DR EMBL: M94128; G192526; JOINED.
DR EMBL: M94127; G192526; JOINED.
DR PIR: A46476; A46476.
DR HSP: P19438; 1TNF.
DR MGD; MGI:88336; CD40.
DR PROSITE: PS00652; TNFR_NGFR_1; 1.
DR PROSITE: PS00500; TNFR_NGFR_2; 4.
KW RECEPTOR; B-CELL; GLYCOPROTEIN; TRANSMEMBRANE; REPEAT; SIGNAL.
FT CHAIN 1 19 POTENTIAL.
FT SIGNAL 1 19 POTENTIAL.
FT DOMAIN 20 289 CD40L RECEPTOR.
FT TRANSMEM 20 193 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 194 215 POTENTIAL.
FT DOMAIN 216 289 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 25 187 4 X TNFR-CYS.
FT REPEAT 25 60 TNFR-CYS 1.
FT REPEAT 61 103 TNFR-CYS 2.
FT REPEAT 104 144 TNFR-CYS 3.
FT REPEAT 145 187 TNFR-CYS 4.
FT CARBOHYD 153 153 POTENTIAL.
SQ SEQUENCE 289 AA; 32111 MW; D8D70A2C CRC32;

Query Match 12.7%; Score 111; DB 1; Length 289;
Best Local Similarity 27.1%; Pred. No. 0.0027;
Matches 42; Conservative 9; Mismatches 58; Indels 46; Gaps 9;

QY 13 LEAAILFLHLACAVSCETGDSGKQEFKDSGNCVLCQKQSGMELSKGFGYGEDAQ 72
DB 10 LWGCLLTVAVHLG---QCVT--CSDKQYIHD--GQC--CDLQPSRLTSHC--TALEKTQ 58
QY 73 CVPCPRHFKEDMG-----FOKCKPCADCALVNRFORANCST--SDAVC----- 116
DB 59 CHPDSEGFSAQWNRREIRCHGRHCEP-----NQLRVKKEGTAEBSDTVCTCKEGQHC 111
QY 116 -----GDCLPGFYRRTKLVGFQDMCEVPC 139
DB 112 TSKDCAACAHPTPCIPFGVMEATETTDVTCCHPC 146

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 1999, 21:39:58 ; Search time 48.52 Seconds

(without alignments)
171.693 Million cell updates/sec

Title: US-09-212-270-4

Sequence: 1 MALKVPLHRYVLEFAILFL.....ODMECVPCGDPPEPPHCE 151

Scoring table: BLOSUM62

Searched: 180763 seqs, 55169189 residues

Database :

SPTREMBL_8:*
1: sp_fungi:*
2: sp_human:*
3: sp_invertebrate:*
4: sp_mammal:*
5: sp_mhc:*
6: sp_organelle:*
7: sp_phage:*
8: sp_plant:*
9: sp_bacteria:*
10: sp_rodent:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|---------------------|
| 1 | 146 | 16.7 | 267 | 4 | 002764 | 002764 oryctolagus |
| 2 | 141.5 | 16.2 | 277 | 2 | 014866 | 014866 homo sapien |
| 3 | 141.5 | 16.2 | 426 | 2 | 014865 | 014865 homo sapien |
| 4 | 136 | 15.6 | 348 | 11 | 057277 | 057277 monkeypox v |
| 5 | 136 | 15.6 | 349 | 11 | 057281 | 057281 monkeypox v |
| 6 | 136 | 15.6 | 349 | 11 | 057099 | 057099 monkeypox v |
| 7 | 136 | 15.6 | 349 | 11 | 057100 | 057100 monkeypox v |
| 8 | 136 | 15.6 | 349 | 11 | 057101 | 057101 monkeypox v |
| 9 | 136 | 15.6 | 349 | 11 | 057102 | 057102 monkeypox v |
| 10 | 136 | 15.6 | 348 | 11 | 057108 | 057108 monkeypox v |
| 11 | 132 | 15.1 | 348 | 11 | 057103 | 057103 monkeypox v |
| 12 | 131 | 15.0 | 186 | 11 | 072735 | 072735 compox viru |
| 13 | 130.5 | 14.9 | 350 | 11 | 057116 | 057116 compox viru |
| 14 | 129 | 14.8 | 625 | 10 | 035305 | 035305 mus musculu |
| 15 | 128.5 | 14.7 | 283 | 2 | 092956 | 092956 homo sapien |
| 16 | 127 | 14.5 | 380 | 2 | 000280 | 000280 homo sapien |
| 17 | 126 | 14.4 | 351 | 11 | 073559 | 073559 compox viru |
| 18 | 125 | 14.3 | 372 | 2 | 000279 | 000279 homo sapien |
| 19 | 124.5 | 14.3 | 355 | 11 | 085308 | 085308 compox viru |
| 20 | 123.5 | 14.1 | 418 | 2 | 000275 | 000275 homo sapien |
| 21 | 123.5 | 14.1 | 253 | 2 | 000276 | 000276 homo sapien |
| 22 | 121 | 13.9 | 351 | 11 | 057117 | 057117 compox viru |
| 23 | 120 | 13.7 | 350 | 11 | 057123 | 057123 compox viru |
| 24 | 119 | 13.6 | 349 | 11 | 057284 | 057284 camelipox vi |
| 25 | 119 | 13.6 | 349 | 11 | 057098 | 057098 camelipox vi |
| 26 | 118.5 | 13.6 | 349 | 11 | 057305 | 057305 compox viru |
| 27 | 117.5 | 13.5 | 348 | 11 | 057112 | 057112 variola vir |
| 28 | 117.5 | 13.5 | 348 | 11 | 085407 | 085407 variola vir |
| 29 | 117 | 13.4 | 334 | 3 | 024403 | 024403 drosophila |

ALIGNMENTS

| | | | | | | |
|----|-------|------|------|----|--------|---------------------|
| 30 | 117 | 13.4 | 349 | 11 | 057110 | 057110 variola vir |
| 31 | 117 | 13.4 | 349 | 11 | 057111 | 057111 variola vir |
| 32 | 117 | 13.4 | 349 | 11 | 089118 | 089118 variola vir |
| 33 | 117 | 13.4 | 349 | 11 | 089098 | 089098 variola vir |
| 34 | 116 | 13.3 | 351 | 11 | 057121 | 057121 compox viru |
| 35 | 113 | 12.9 | 401 | 10 | 070202 | 070202 mus musculu |
| 36 | 113 | 12.9 | 349 | 11 | 057097 | 057097 camelipox vi |
| 37 | 111 | 12.7 | 1737 | 2 | 075097 | 075097 homo sapien |
| 38 | 110.5 | 12.7 | 326 | 11 | 057120 | 057120 compox viru |
| 39 | 110.5 | 12.7 | 326 | 11 | 057122 | 057122 compox viru |
| 40 | 110 | 12.6 | 347 | 11 | 057115 | 057115 compox viru |
| 41 | 109 | 12.5 | 265 | 3 | 024404 | 024404 drosophila |
| 42 | 109 | 12.5 | 349 | 11 | 057109 | 057109 variola vir |
| 43 | 109 | 12.5 | 360 | 11 | 057118 | 057118 compox viru |
| 44 | 109 | 12.5 | 347 | 11 | 057119 | 057119 compox viru |
| 45 | 108 | 12.4 | 401 | 10 | 008712 | 008712 mus musculu |

| | | | | | | |
|--|--------------|-----------|-----------------|--|--|--|
| RESULT 1 | | | | | | |
| ID 002764 | PRELIMINARY: | PRT: | 267 AA. | | | |
| AC 002764 | | | | | | |
| DT 01-JUL-1997 (TREMBLREL. 04, CREATED) | | | | | | |
| DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE) | | | | | | |
| DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE) | | | | | | |
| DE OX40 PRECURSOR (FRAGMENT) | | | | | | |
| OS ORYCTOLAGUS CUNICULUS (RABBIT) | | | | | | |
| OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; | | | | | | |
| OC LACOMORPHA; LEPORIDAE; ORYCTOLAGUS. | | | | | | |
| RN [1] | | | | | | |
| RP SEQUENCE FROM N.A. | | | | | | |
| RC STRAIN-CHB:HM; | | | | | | |
| RA ISONO T., SEPTO A.; | | | | | | |
| RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS. | | | | | | |
| DR EMBL; AB003911; D1020869; . | | | | | | |
| DR PROSITE; PS00652; TNFR_NGFR_1; 2. | | | | | | |
| DR PFM; PF00020; TNFR_C6; 3. | | | | | | |
| KW SIGNAL. | | | | | | |
| FT NON_TER | 1 | 1 | | | | |
| FT SIGNAL | <1 | 18 | POTENTIAL. | | | |
| FT CHAIN | 19 | 267 | OX40. | | | |
| SQ SEQUENCE | 267 AA; | 28489 MW; | FB3COC5E CRC32; | | | |

Query Match 16.7%; Score 146; DB 4; Length 267;
Best Local Similarity 34.7%; Pred. No. 2.7e-05;
Matches 33; Conservative 11; Mismatches 37; Indels 14; Gaps 6;

| | | |
|--|---|--------------|
| QY 50 | CKOCPGKELKECFGEADQCVPCRPHEKEDMGFKCPACADLVNRFQ---- | RA 105 |
| DB 37 | CLECPGIGMVSRC--NSQDICHPCPEGFYNEVNTQACKPCTQC--NRRSGSEFQ 91 | |
| QY 106 | NCSHTSDAVCGDLPFGYKRTLVGFQ--DMECVPC 139 | |
| DB 92 | ECTHRTDYVC--RCRPG--TQPLNKGKHYGVCAPC 122 | |
| RESULT 2 | | |
| ID 014866 | PRELIMINARY: | PRT: 277 AA. |
| AC 014866 | | |
| DT 01-JAN-1998 (TREMBLREL. 05, CREATED) | | |
| DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE) | | |
| DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE) | | |
| DE SOUBLE DEATH RECEPTOR 3 BETA. | | |
| GN DR3. | | |
| OS HOMO SAPIENS (HUMAN). | | |
| OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES; | | |
| OC CATARRHINI; HOMINIDAE; HOMO. | | |
| RN [1] | | |

RP SEQUENCE FROM N.A.
 RA MARCOCHA K., RIBEIRO P., RENARD N., CHARLOT C., COIFFIER B.,
 RA SALES G.;
 RL SUBMITTED (SEP-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: AF026071; G2570833;
 DR PROSITE: PS00652; TNR_NGFR_1; 2.
 DR PFAM: PF00020; TNR_C6; 2.
 DR SEQUENCE 277 AA; 29111 MW; 06E165C7 CRC32;

Query Match 16.2%; Score 141.5; DB 2; Length 277;
 Best Local Similarity 33.1%; Pred. No. 9.4e-05;
 Matches 40; Conservative 12; Mismatches 36; Indels 33; Gaps 8;

QY 33 DCSRQGEFFKDRSG-----NC--VLCKQCG--PGMELSKGCGFGYGEDACQVPCRPFRHFE 83
 Db 88 ECARQACDEQASQVALENCSAVALDTRCGCKPGMEV--EC-----QVSCQVSSSPF--- 137
 QY 84 DMGFOCKPCADCALVNRFORANCSTHTSDAVCGDCLPGFYRKTKLVGFQDMCEVPCGDDP 143
 Db 137 -----YCQPCLDGALHRRHRLCLSR-RDTDCGTCLPGFYE-----HGDGCVSCPTTP 183
 QY 144 P 144
 Db 184 P 184

RESULT 3
 014865 PRELIMINARY; PRT; 426 AA.
 ID 014865;
 AC 014865;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE DEATH RECEPTOR 3 BETA.
 DR3.
 OS HOMO SAPIENS (HUMAN).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 OC CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 98113360.
 RA MARCOCHA K., RIBEIRO P., CHARLOT C., RENARD N., COIFFIER B.,
 RA SALES G.;
 RL "A new death receptor 3 isoform: expression in human lymphoid cell
 lines and non-Hodgkin's lymphomas";
 RL BIOCHEM. BIOPHYS. RES. COMMUN. 242:376-379(1998).
 DR EMBL: AF026070; G2570831;
 DR PROSITE: PS00652; TNR_NGFR_1; 2.
 DR PFAM: PF00020; TNR_C6; 2.
 DR PFAM: PF00531; death; 1.
 DR SEQUENCE 426 AA; 45950 MW; 87E00821 CRC32;

Query Match 16.2%; Score 141.5; DB 2; Length 426;
 Best Local Similarity 33.1%; Pred. No. 9.4e-05;
 Matches 40; Conservative 12; Mismatches 36; Indels 33; Gaps 8;

QY 33 DCSRQGEFFKDRSG-----NC--VLCKQCG--PGMELSKGCGFGYGEDACQVPCRPFRHFE 83
 Db 88 ECARQACDEQASQVALENCSAVALDTRCGCKPGMEV--EC-----QVSCQVSSSPF--- 137
 QY 84 DMGFOCKPCADCALVNRFORANCSTHTSDAVCGDCLPGFYRKTKLVGFQDMCEVPCGDDP 143
 Db 137 -----YCQPCLDGALHRRHRLCLSR-RDTDCGTCLPGFYE-----HGDGCVSCPTTP 183
 QY 144 P 144
 Db 184 P 184

RESULT 4
 057277

ID 057277 PRELIMINARY; PRT; 348 AA.
 AC 057277;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYVIRIDAE; CHORDOPOXYVIRINAE;
 OC ORTHOPOXYVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ZAIRE-1996(96-17), ZAIRE-1996 (96-16);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U88543; G2738061;
 DR EMBL: U87841; G2738061;
 DR PROSITE: PS00652; TNR_NGFR_1; 2.
 DR SEQUENCE 348 AA; 38212 MW; 1AAAB69 CRC32;

Query Match 15.6%; Score 136; DB 11; Length 348;
 Best Local Similarity 26.3%; Pred. No. 0.00022;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RYVLFALLFLHLACKVSC-----ETGDCSRQGEFFKDRSGNCVLCRQCGPG 56
 Db 2 RSYLYSYILFL-----SCIINGRDIALPHAPSNGKC-KDNEYRSRN---LCGICSPCG 50
 QY 57 MELSKGCGFGYGEDACVCRPHRF-KEDWGFOCKPCADCALVNRFORANCSTHTSDAVC 115
 Db 51 TYASRLC--DSKTNQCTPCGSDFTSHNNHLQACLSCGRDSDVVERRSNTTHNRIC 108
 QY 116 GDCLPGFY-----RKTKL-VGF-----QDMCEVPCG 140
 Db 109 -ECSPGYCLLKGSSGCRFCISKTKGIGYGSYTSSTDVLCSPCG 154

RESULT 5
 057291 PRELIMINARY; PRT; 349 AA.
 ID 057291;
 AC 057291;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYVIRIDAE; CHORDOPOXYVIRINAE;
 OC ORTHOPOXYVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-VARIOUS STRAINS;
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U88144; G2738087;
 DR EMBL: U87842; G2738063;
 DR EMBL: U87994; G2738079;
 DR EMBL: U87995; G2738081;
 DR EMBL: U88143; G2738085;
 DR PROSITE: PS00652; TNR_NGFR_1; 2.
 DR SEQUENCE 349 AA; 38295 MW; 7313FCF9 CRC32;

Query Match 15.6%; Score 136; DB 11; Length 349;
 Best Local Similarity 26.3%; Pred. No. 0.00023;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RYVLFALLFLHLACKVSC-----ETGDCSRQGEFFKDRSGNCVLCRQCGPG 56
 Db 2 RSYLYSYILFL-----SCIINGRDIALPHAPSNGKC-KDNEYRSRN---LCGICSPCG 50
 QY 57 MELSKGCGFGYGEDACVCRPHRF-KEDWGFOCKPCADCALVNRFORANCSTHTSDAVC 115

Db 51 TYASRLC--DSKNTQCTCGSDTFTSHNNHLQACLSGRCDSNQVETRSCWTTNRLC 108
 QY 116 GDCLPGEFY-----RKTKL-VGF-----ODMECVPCG 140
 Db 109 -ECSPGYCLLKGASGCRCTCISRTKCGIGVSGYSTGDIYCSPCG 154

RESULT 6

057099 PRELIMINARY: PRT: 349 AA.
 AC 057099;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-SIERRA LEONE-1970 (70-0266);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U87843: G2738065; -;
 DR PROSITE: P500652; TNFR_NGFR_1; 2.
 SQ SEQUENCE 349 AA; 38321 MW; 5315315C CRC32;

Query Match

Best Local Similarity 15.6%; Score 136; DB 11; Length 349;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RTVLFAAILFLHLACKVSC-----ETGDCSRQOEFKDRSGNCVLCCKCGPG 56
 Db 2 RSLVLYIILFL-----SCIINGRDIAPHAPSNGKC-KDNEYRSRN---LCCLSCPPG 50
 QY 57 MELSKCGGFGYGEDAQVCPRHFR-KEDMGFOKCKPCADCALVNFQANCSHTSDAYC 115
 Db 51 TYASRLC--DSKNTQCTCGSDTFTSHNNHLQACLSGRCDSNQVETRSCWTTNRLC 108
 QY 116 GDCLPGEFY-----RKTKL-VGF-----ODMECVPCG 140
 Db 109 -ECSPGYCLLKGALGCRCTCISRTKCGIGVSGYSTGDIYCSPCG 154

RESULT 7

057100 PRELIMINARY: PRT: 349 AA.
 AC 057100;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-NIGERIA-1971 (71-0082);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U87844: G2738067; -;
 DR PROSITE: P500652; TNFR_NGFR_1; 2.
 SQ SEQUENCE 349 AA; 38239 MW; F8871DD2 CRC32;

Query Match 15.6%; Score 136; DB 11; Length 349;
 Best Local Similarity 26.3%; Pred. No. 0.00023;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RTVLFAAILFLHLACKVSC-----ETGDCSRQOEFKDRSGNCVLCCKCGPG 56
 Db 2 RSLVLYIILFL-----SCIINGRDIAPHAPSNGKC-KDNEYRSRN---LCCLSCPPG 50
 QY 57 MELSKCGGFGYGEDAQVCPRHFR-KEDMGFOKCKPCADCALVNFQANCSHTSDAYC 115
 Db 51 TYASRLC--DSKNTQCTCGSDTFTSHNNHLQACLSGRCDSNQVETRSCWTTNRLC 108

QY 116 GDCLPGEFY-----RKTKL-VGF-----ODMECVPCG 140
 Db 109 -ECSPGYCLLKGASGCRCTCISRTKCGIGVSGYSTGDIYCSPCG 154

RESULT 8

057101 PRELIMINARY: PRT: 349 AA.
 AC 057101;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-FAIRE-1977 (77-0666);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U87845: G2738069; -;
 DR PROSITE: P500652; TNFR_NGFR_1; 2.
 SQ SEQUENCE 349 AA; 38311 MW; 4FAF8AF1 CRC32;

Query Match

Best Local Similarity 15.6%; Score 136; DB 11; Length 349;
 Matches 44; Conservative 22; Mismatches 51; Indels 50; Gaps 10;

QY 10 RTVLFAAILFLHLACKVSC-----ETGDCSRQOEFKDRSGNCVLCCKCGPG 56
 Db 2 RSLVLYIILFL-----SCIINGRDIAPHAPSNGKC-KDNEYRSRN---LCCLSCPPG 50
 QY 57 MELSKCGGFGYGEDAQVCPRHFR-KEDMGFOKCKPCADCALVNFQANCSHTSDAYC 115
 Db 51 TYASRLC--DSKNTQCTCGSDTFTSHNNHLQACLSGRCDSNQVETRSCWTTNRLC 108
 QY 116 GDCLPGEFY-----RKTKL-VGF-----ODMECVPCG 140
 Db 109 -ECSPGYCLLKGSSGCRCTCISRTKCGIGVSGYSTGDIYCSPCG 154

RESULT 9

057102 PRELIMINARY: PRT: 349 AA.
 AC 057102;
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
 DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRMB.
 OS MONKEYPOX VIRUS.
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXVIRIDAE; CHORDOPOXVIRINAE;
 OC ORTHOPOXVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BENIN-1978 (78-3945);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U87846: G2738071; -;
 DR PROSITE: P500652; TNFR_NGFR_1; 2.
 SQ SEQUENCE 349 AA; 38308 MW; 17182545 CRC32;

DT 01-AUG-1998 (TREMBLREL. 07, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR RECEPTOR II HOMOLOG.
 GN CRM
 OS COMPOX VIRUS (CPV).
 OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; POXYIRIDAE; CHORDOPOXYIRINAE;
 OC ORTHOPOXYVIRUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MUNICH OPV 89/2(CAT);
 RA LOPAREV V.N., PARSONS J.M., ESPOSITO J.J.;
 RL SUBMITTED (FEB-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
 DR EMBL: U90226; G2738129; -
 SQ SEQUENCE 350 AA; 38216 MW; 5FF8BE39 CRC32;

Query Match 14.8%; Score 130.5; DB 11; Length 350;
 Best Local Similarity 25.9%; Pred. No. 0.00064;
 Matches 43; Conservative 21; Mismatches 55; Indels 47; Gaps 10;

QY 10 RTVFAILFLHLACKVSC-----ETGDCSRQGEKDKSGNCVLCCKGCGPG 56
 DB 2 KSVLYSTILFL-----SCIINGRVAPYAPSPNGRC-KDNEY-NRHNLC--CLSCPPG 50
 QY 57 MELSKGCGFGEDACVPCPRHFR-KEDMGFCCKPCADALVNRFORANCSTSDAVC 115
 DB 51 TYASRLCDSTNTNTCTPCGSDFTSRNNHLPACLSGNGRCDSDNOVETSCNTTHNRIC 110
 QY 116 GDCLPGFY-----RKTKL-VGF-----QDMCEVPCG 140
 DB 111 -DCAPGYCYLLKSGKACVSGTCGIGYVSGHTPTGDIVICPCG 155

RESULT 14
 ID 035305 PRELIMINARY; PRT; 625 AA.
 AC 035305;
 DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
 DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE RECEPTOR ACTIVATOR OF NF-KAPPAB (RANK).
 GN RANK.
 OS MUS MUSCULUS (MOUSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
 OC SCIUROGNATHI; MURIDAE; MURINAE; MUS.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-LIVER EPITHELIUM;
 RX MEDLINE: 98032977.
 RA ANDERSON D.M., MARASKOVSKY E., BILLINGSLEY W.L., DOUGALL W.C.,
 RA TOMERSKO M.E., ROUX E.R., TEEPE M.C., DOBOSE R.F., COSMAN D.,
 RA GALIBERT L.;
 RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
 RT and dendritic-cell function.";
 RL NATURE 390:175-179(1997).
 DR EMBL: AF019046; G2612920; -
 DR MGD: MGI1194917; RANK.
 DR PRAM: PF00020; TNFR_C6; 3.
 DR SEQUENCE 625 AA; 66621 MW; A01E6C5B CRC32;

Query Match 14.8%; Score 129; DB 10; Length 625;
 Best Local Similarity 29.9%; Pred. No. 0.0014;
 Matches 44; Conservative 20; Mismatches 57; Indels 26; Gaps 10;

QY 13 LFAAILFLHLACKVSCETGDCSRQGEKDKSGNCVLCCKGCGPGMELSKGCGFGYGE 72
 DB 15 LLAICVLLVPLQVLYQV-TTPCTOBRNY-EHLGRK--CSRCEPKYILSSKC--TPTSDSV 68
 QY 73 CVPCPRHFRKEDMGFQ-KC--KPC-ADCALV-----NRFORANCSTSDAVC- 116
 DB 69 CLPCGPDEYLDITWNEEDKCLHLKVCDAAGKALVADPGNHTAPRRCACTAGYHWNMSDCECC 128
 QY 116 --GDCLPGFYRKTLYGFDMECVPC 139

DB 129 RRTNECAPGTGAQHPLQLNNDYCTPC 155
 :| ||| :| :| |||
 RESULT 15
 ID 092956 PRELIMINARY; PRT; 283 AA.
 AC 092956;
 DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
 DT 01-JUL-1997 (TREMBLREL. 04, LAST SEQUENCE UPDATE)
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
 DE HERPESVIRUS ENTRY MEDIATOR.
 OS HOMO SAPIENS (HUMAN)
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
 OC CATARRHINI; HOMINIDAE; HOMO.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 97053782.
 RA MONTOMERY R.I., WARNER M.S., LUM B.J., SPEAR P.G.;
 RT "Herpes simplex virus-1 entry into cells mediated by a novel member of
 RT the TNF/NGF receptor family.";
 RL CELL 87:427-436(1996).
 DR EMBL: U70321; E319244; -
 DR PROSITE: P500652; TNFR_NGFR_1; 1.
 DR PRAM: PF00020; TNFR_C6; 2.
 DR SEQUENCE 283 AA; 30420 MW; B67FBE7B CRC32;

Query Match 14.7%; Score 128.5; DB 2; Length 283;
 Best Local Similarity 27.7%; Pred. No. 0.00079;
 Matches 46; Conservative 18; Mismatches 67; Indels 35; Gaps 10;

QY 10 RTVFAILFLHLACKVSCETGDCSRQGEKDKSGNCVLCCKGCGPGMELSKGCGFGYGE 69
 DB 17 RTDYLRLVLYLFLGAPCYAPALPSCKEDEYPIVG-EC--CPKSPCYRYKACGELTG- 73
 QY 70 DAQCVPCPRHFRKEDW-GFQCKKPCADC--ALVNRFORANCSTSDAVCGDCLPGFY--- 124
 DB 73 -TYCEPCPPGYIAHLNGLKLCLOCCQCDPAMGIRASR-NCSRTENAVCG-CSPGHFCIV 129
 QY 124 -----RKTLYG--FQDMCEVPCGDPPEPEPH 149
 DB 130 QDDHCAACATATSSPGQYQKGGTESQDTLQNC--PPTGISPN 173

Search completed: May 14, 1999, 21:39:58
 Job time: 7992 sec

This Page Blank (uspto)

Run on: May 13, 1999, 17:21:19 ; Search time 42.79 Seconds

(without alignments)
96.896 Million cell updates/sec

```

Title: US-09-212-270-5
Perfect score: 1061
Sequence: 1 MVLQTQTELOS LRREVSRL.....RARA KLNI SPHGTELG FVKL 205

```

Scoring table: BLOSUM62

Searched: 162890 seqs, 20225328 residues

Database : A_Geneseq_34:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|---------------------|
| 1 | 1019 | 96.0 | 233 | 1 | W37002 | Human tumour necro |
| 2 | 836 | 78.8 | 168 | 1 | W37003 | Human tumour necro |
| 3 | 229.5 | 21.6 | 285 | 1 | W58391 | Homo sapiens neurit |
| 4 | 229.5 | 21.6 | 285 | 1 | W62461 | Human T cell surfat |
| 5 | 229.5 | 21.6 | 285 | 1 | W73043 | Tumour necrosis fa |
| 6 | 227.5 | 21.4 | 285 | 1 | W62460 | Human T cell surfat |
| 7 | 217 | 20.5 | 266 | 1 | W62462 | Human T cell surfat |
| 8 | 197 | 18.6 | 150 | 1 | W62463 | Mouse T cell surfat |
| 9 | 99.5 | 9.4 | 155 | 1 | R44145 | TNF RGD mutein #1. |
| 10 | 98.5 | 9.3 | 158 | 1 | P60556 | Sequence of mature |
| 11 | 98.5 | 9.3 | 161 | 1 | R44151 | TNF RGD mutein #7. |
| 12 | 98.5 | 9.3 | 163 | 1 | R44152 | TNF RGD mutein #8. |
| 13 | 97.5 | 9.2 | 145 | 1 | P82043 | Sequence of new tu |
| 14 | 97.5 | 9.2 | 158 | 1 | R44148 | TNF RGD mutein #4. |
| 15 | 96.5 | 9.1 | 155 | 1 | R05174 | Tumoricidal polypee |
| 16 | 96.5 | 9.1 | 157 | 1 | R14893 | Antitumoural polyp |
| 17 | 96.2 | 9.1 | 216 | 1 | W11702 | Alpha-1-thymosin-a |
| 18 | 96 | 9.0 | 235 | 1 | R05328 | Rat tumour necrosi |
| 19 | 96 | 9.0 | 235 | 1 | R05396 | Rat tumour necrosi |
| 20 | 95.5 | 9.0 | 169 | 1 | P82606 | Fusion protein der |
| 21 | 95.5 | 9.0 | 157 | 1 | R14892 | Antitumoural polyp |
| 22 | 95.5 | 9.0 | 155 | 1 | R44146 | TNF RGD mutein #2. |
| 23 | 95.5 | 9.0 | 157 | 1 | R62472 | Tumour necrosis fa |
| 24 | 94.5 | 8.9 | 158 | 1 | P70636 | Sequence of tumour |
| 25 | 94.5 | 8.9 | 151 | 1 | R11615 | Tumour Necrosis fa |
| 26 | 94.5 | 8.9 | 151 | 1 | R11968 | Tumour Necrosis fa |
| 27 | 94.5 | 8.9 | 155 | 1 | R22314 | TNF polypeptide mu |
| 28 | 94.5 | 8.9 | 162 | 1 | W48650 | Heavy chain of huma |
| 29 | 93.5 | 8.8 | 171 | 1 | P82607 | Fusion protein der |
| 30 | 93.5 | 8.8 | 176 | 1 | P82608 | Fusion protein der |
| 31 | 93.5 | 8.8 | 150 | 1 | R08151 | Polypeptide with a |
| 32 | 93.5 | 8.8 | 157 | 1 | P61653 | L-M cell prolifera |
| 33 | 93.5 | 8.8 | 157 | 1 | R14800 | Antitumoural polyp |
| 34 | 93.5 | 8.8 | 155 | 1 | R22313 | TNF polypeptide mu |
| 35 | 93.5 | 8.8 | 150 | 1 | R29609 | TNF analogue #2. P |
| 36 | 93.5 | 8.8 | 155 | 1 | R44144 | Wild type TNF. Tum |
| 37 | 93.5 | 8.8 | 158 | 1 | R44149 | TNF RGD mutein #5. |
| 38 | 93.5 | 8.8 | 158 | 1 | R44150 | TNF RGD mutein #6. |
| 39 | 93.5 | 8.8 | 157 | 1 | R62478 | Tumour necrosis fa |
| 40 | 93.5 | 8.8 | 157 | 1 | R62479 | Tumour necrosis fa |
| 41 | 93.5 | 8.8 | 157 | 1 | R63097 | Tumour necrosis fa |
| 42 | 93 | 8.8 | 154 | 1 | R23312 | TNF polypeptide mu |
| 43 | 92.5 | 8.7 | 151 | 1 | R15816 | TNF analogue encod |

| | 92.5 | 8.7 | 151 | 1 | R14305 | TNF analogue encod |
|----|------|-----|-----|---|--------|--------------------|
| | 92.5 | 8.7 | 157 | 1 | R20627 | Synthetic hTNF mut |
| 44 | 92.5 | 8.7 | 151 | 1 | R14305 | TNF analogue encod |
| 45 | 92.5 | 8.7 | 157 | 1 | R20627 | Synthetic hTNF mut |

ALIGNMENTS

```

RESULT      1
ID          W37002 standard; Protein; 233 AA.
AC          W37002;
DT          14-APR-1998 (first entry)
DE          Human tumour necrosis factor delta.
KW          Human; tumour necrosis factor delta; TNF-delta; TNF-epsilon; cancer;
KW          tumour necrosis factor epsilon; T cell proliferation;
KW          immune regulation; inflammatory response.
OS          Homo sapiens.
FH          key
FT          Location/Qualifiers
FT          Misc_difference 231..233
FT          /note= "Not given in the specification but present
FT          if the cDNA is fully decoded"
PN          M09733902-A1.
PD          18-SEP-1997.
PE          14-MAR-1996; U03774.
PR          14-MAR-1996; WO-U03774.
PR          (HUMA-) HUMAN GENOME SCI INC.
PI          Gentz RL, Na J, Yu G;
PI          WPI: 97-470810/43.
DR          N-PSDB: V00493.
PT          Nucleic acid encoding human tumour necrosis factors delta and
PT          epsilon - useful for destroying cancer cells, or mediating T cell
PT          proliferation, immune regulation and inflammatory responses etc.
PS          Claim 18: Fig 1: 83pp: English.
CC          The present sequence represents human tumour necrosis factor (TNF)
CC          delta. The TNF can be used to treat patients deficient in this factor
CC          (optionally by in vivo expression). It is a ligand for TNF and can
CC          destroy (by inducing apoptosis) some transformed cell lines (for tumour
CC          treatment); mediate cell activity and proliferation (including T cells
CC          to stimulate an immune response to viral, bacterial or parasitic
CC          infections), also to eliminate autoreactive T cells in e.g. type I
CC          diabetes), and is functionally linked as a primary mediator of immune
CC          regulation and the inflammatory response. A typical application of the
CC          protein, its agonists or antagonists is prevention of septic shock,
CC          inflammation, cerebral malaria, human immunodeficiency virus activation,
CC          graft-host reaction, bone resorption (osteoporosis), rheumatoid
CC          arthritis and cachexia, also to promote wound healing (by upregulating
CC          cell adhesion) and to regulate haematopoiesis. The protein can also be
CC          used to isolate its cognate receptors (or receptor-encoding genes).
CC          The cDNA can be used to produce recombinant proteins, and fragments of
CC          it to isolate full-length or related sequences, for diagnostic detection
CC          of mutations (indicative of disease or susceptibility) and for
CC          chromosome identification. Analysis of patient samples for presence of
CC          the new protein (e.g. using specific antibodies) can be used
CC          diagnostically.
SQ          Sequence      233 AA;

```

| | Query Match | 96.0% | Score 1019; | DB 1; | Length 233; |
|----|---|-----------------|--------------------|-----------|-------------|
| | Best Local Similarity | 97.1%; | Pred. No. 1.7e-99; | | |
| | Matches 199; | Conservative 1; | Mismatches 5; | Indels 0; | Gaps 0; |
| QY | 1 MWLTQQTLELQSLKREVSRLOQTGGSPSQNGEYPMQSLPEQSSDALLEAMENERSKKRA | 60 | | | |
| | | | | | |
| Db | 29 MALLQQTLELQSLKREVSRLOQTGGSPSQNGEYPMQSLPEQSSDALLEAMENERSKKRA | 88 | | | |
| | | | | | |
| QY | 61 VLTQKQKQKQSHVLHPINATSKDSDVTEVMQPALRRGRLOAQGYVRIQDAGVYLL | 120 | | | |
| | | | | | |
| Db | 89 VLSQKQKQKQSHVLHPINATSKDSDVTEVMQPALRRGRLOAQGYVRIQDAGVYLL | 148 | | | |
| | | | | | |
| QY | 121 YSQVLFQVYTTMGVVSREGGQRETLFRCIRSNPSPHPDRAYNSQYISAGVPHLQGDLL | 180 | | | |
| | | | | | |
| Db | 149 YSQVLFQVYTTMGVVSREGGQRETLFRCIRSNPSPHPDRAYNSQYISAGVPHLQGDILL | 208 | | | |

QY 181 SVIIPRAKLNLSPHGTFLGFVKL 205
 ||||||||||||||||||
 Db 209 SVIIPRAKLNLSPHGTFLGFVKL 233

RESULT 2

W37003
 ID M37003 standard; Protein: 168 AA.
 AC W37003;
 DT 14-APR-1998 (first entry)
 DE Human tumour necrosis factor epsilon.
 KW Human tumour necrosis factor delta; TNF-delta; TNF-epsilon; cancer;
 KM tumour necrosis factor epsilon; T cell proliferation;
 OS Homo sapiens.
 PN MO9733902-A1.
 PD 18-SEP-1997.
 PR 14-MAR-1996; WO-003774.
 PR (HUMA-) HUMAN GENOME SCI INC.
 PA Gentz RL, Ni J, Yu G;
 PI MPI: 97-470810/43.
 DR N-PSDB; V00494.
 PT Nucleic acid encoding human tumour necrosis factors delta and epsilon - useful for destroying cancer cells, or mediating T cell proliferation, immune regulation and inflammatory responses etc.
 PS Claim 20; Fig 2; 83pp; English.
 CC The present sequence represents human tumour necrosis factor (TNF) epsilon. The TNF can be used to treat patients deficient in this factor (optionally by in vivo expression). It is a ligand for TNF and can destroy (by inducing apoptosis) some transformed cell lines (for tumour treatment); mediate cell activity and proliferation (including T cells to stimulate an immune response to viral, bacterial or parasitic infections, also to eliminate autoreactive T cells in e.g. type I diabetes), and is functionally linked as a primary mediator of immune regulation and the inflammatory response. A typical application of the protein, its agonists or antagonists is prevention of septic shock, inflammation, cerebral malaria, human immunodeficiency virus activation, graft-host reaction, bone resorption (osteoporosis), rheumatoid arthritis and cachexia, also to promote wound healing (by upregulating cell adhesion) and to regulate haematopoiesis. The protein can also be used to isolate its cognate receptors (or receptor-encoding genes).
 CC The CDNA can be used to produce recombinant proteins, and fragments of it to isolate full-length or related sequences, for diagnostic detection of mutations (indicative of disease or susceptibility) and for chromosome identification. Analysis of patient samples for presence of the new protein (e.g. using specific antibodies) can be used diagnostically.
 CC Sequence 168 AA;
 SQ

Query Match 78.8%; Score 836; DB 1; Length 168;
 Best Local Similarity 89.1%; Pred. No. 1,9e-80;
 Matches 164; Conservative 2; Mismatches 2; Indels 16; Gaps 1;

QY 22 GTGSPONGEGYPMOSLPQSSDALTEWNGERSRKRRAVLTKOKKHSLVHLVPINAT 81
 ||||||||||||||||||
 Db 1 GTGSPONGEGYPMOSLPQSSDALTEWNGERSRKRRAVLTKOKKHSLVHLVPINAT 46
 QY 82 SKDDSDVTEVMOPALRRGGLAOGYGVRIODAGVYLLYSQVLFQDVFPTMGQVVSREG 141
 ||||||||||||||||||
 Db 46 -KNDSDVTEVMOPALRRGGLAOGYGVRIODAGVYLLYSQVLFQDVFPTMGQVVSREG 104
 QY 142 QGQOETLFRICRSMPSHPDAVNSQVSAGVPHLHQGDILSVIIPRAKLNLSPHGTFLG 201
 ||||||||||||||||||
 Db 105 QGQOETLFRICRSMPSHPDAVNSQVSAGVPHLHQGDILSVIIPRAKLNLSPHGTFLG 164
 QY 202 FVKL 205
 |||||
 Db 165 FVKL 168

RESULT 3

W58391
 ID M58391 standard; Protein: 285 AA.
 AC W58391;
 DT 11-SEP-1998 (first entry)
 DE Homo sapiens neutrokin alpha protein.
 KW neutrokin alpha; cell proliferation; differentiation; migration;
 KM cytotoxicity; cell death; treatment; tumour; infection; inflammation;
 KW wound healing; immunodeficiency; autoimmune disease; graft rejection;
 KM fibrotic disorder; haematopoiesis; sepsis; shock; malaria; HIV; AIDS;
 KW acquired immune deficiency syndrome; rheumatoid arthritis; silicosis;
 OS Homo sapiens.
 FH Key
 FT Domain
 FT Location/Qualifiers
 FT 1..46 "intracellular domain"
 FT 47..72 "note="transmembrane domain"
 FT 73..285 "note="extracellular domain"
 FT Domain
 FT MO9818921-A1.
 PD 07-MAY-1998.
 PR 25-OCT-1996; U17957.
 PR 25-OCT-1996; WO-U17957.
 PA (HUMA-) HUMAN GENOME SCI INC.
 PI Ebner R, Ni J, Yu G;
 PI MPI: 98-272216/24.
 DR N-PSDB; V30934.
 PT New isolated human Neutrokin alpha - used to develop products for diagnosis and treatment of e.g. tumours, infections,
 PT immunodeficiencies or autoimmune diseases
 PS Claim 17; Fig 1; 104pp; English.
 CC The sequence is that of the neutrokin alpha protein.
 CC Neutrokin alpha (NA) polypeptides modulate cell proliferation, differentiation, migration, cytotoxicity and cell death.
 CC They can be used to treat e.g. tumour and tumour metastasis, infections by bacteria, viruses and other parasites, immunodeficiencies, inflammatory diseases, lymphadenopathy, autoimmune diseases, graft versus host disease and to stimulate peripheral tolerance, destroy some transformed cell lines, mediate cell activation and proliferation, and are functionally linked as primary mediators of immune regulation and inflammatory responses. Such activity is useful for immune enhancement or suppression, myeloprotection, stem cell mobilisation, acute and chronic inflammatory control and treatment of leukaemia. They can also be used to stimulate wound healing and to treat fibrotic disorders including liver cirrhosis, osteoarthritis and pulmonary fibrosis. They can also be used to regulate haematopoiesis, by regulating the activation and differentiation of various haematopoietic progenitor cells, e.g. to release mature leukocytes from the bone marrow following chemotherapy, and in stem cell mobilisation. NA may also be used to treat sepsis, NA antagonists can be used to prevent septic shock, inflammation, cerebral malaria, activation of the HIV virus, graft-host rejection, bone resorption, rheumatoid arthritis and cachexia (wasting or malnutrition).
 CC They can also be used to treat e.g. autoimmune diseases such as multiple sclerosis and insulin-dependent diabetes and inflammatory and infectious diseases such as silicosis, and sarcoidosis, idiopathic pulmonary fibrosis, idiopathic hyper-eosinophilic syndrome, endotoxin shock, atherosclerosis, histamine-mediated allergic reactions and immunological disorders including late phase allergic reactions, chronic urticaria, and atopic dermatitis by inhibiting chemokine-induced mast cell and basophil degranulation and release of histamine. TGF-mediated allergic reactions such as allergic asthma, rhinitis and eczema, inflammatory pulmonary diseases, rheumatoid arthritis, inflammation, degenerative and inflammatory arthropathies, aplastic anaemia, myelodysplastic syndrome, subepithelial basement membrane fibrosis or adult respiratory distress syndrome. The products can also be used for detection, diagnosis and drug screening.
 CC Sequence 285 AA;
 SQ

Query Match 21.6%; Score 229.5; DB 1; Length 285;
 Best Local Similarity 30.9%; Pred. No. 2e-16;
 Matches 67; Conservative 41; Mismatches 86; Indels 23; Gaps 7;

```
QY 7 QTEQLSRREYSRLQGTGSPQNGEGYPMQSLPEQS--SDALEME-----NGESR 56
D 73 QGDLASLRAELQGHHAELPA--GAGAPKAGLEBPAPVATGAKLFEPAPPEGNSONS 130
QY 57 KRAVLTOKOKKSHVHLVPIINAT--SKDDSDVTEVMQPALRGRLQAGYGVRIODA 115
D 131 NKRAVQGEETVTDCCQLINDSETPTIQKSYTFVPMULSFKGSALAEKENILVYKET 190
QY 116 GYVLLYSQVLEFQDVTFTMGQVVSRE-----GQGRQETLFRICISMP-SHPDRAYNSQYS 168
D 191 GYFPTIGQVLYTDKTYAMGHILQKRVHVEFGDELSVTLFRICIONMPEPLPN---NSCYS 247
QY 169 AGVPHLHQGDLISVTPRRAKLNLSPHGTFGLGVKL 205
D 248 AGIAKLEGGDELQALAIPIRENNQISLDGDVTFFGALKL 284

RESULT 4
M62461 standard: Protein: 285 AA.
AC M62461:
DE 05-OCT-1998 (first entry)
KW Human T cell surface antigen 63954 protein sequence #2.
KW Human: 63954; primate; rodent; mouse; T cell surface antigen; mammal;
KW diagnosis; antigen-specific proliferation; cytokine production;
KW immune response; autoimmune disorder; rheumatoid arthritis;
KW systemic lupus erythematosus; Hashimoto's autoimmune thyroiditis.
OS Homo sapiens.
PN M09827114-A2.
PD 25-JUN-1998
PF 16-DEC-1997: U23321.
PR 17-DEC-1996: US-033601.
PA (SCHE ) SCHERING CORP.
PI Gorman DM.
DR WPI: 98-362719/31.
DR N-PSDB: V39985.
PT New isolated polypeptide, 63954 - used to develop products for
PT treating e.g. autoimmune disorders, inflammation, tissue rejection,
PT cancer or degenerative conditions.
PS Claim 1; Page 60-61; 69pp; English.
CC The present sequence is a human T cell surface antigen, designated
CC 63954. The novel protein designated 63954 is expressed on T cells.
CC Protein 63954 can modulate antigen-specific proliferation and cytokine
CC production on effector cells and may potentiate immune cell expansion or
CC apoptosis. 63954 agonists or antagonists may also act as a co-stimulatory
CC molecule for regulation of T cell mediated cell activation, and may cause
CC a shift of T helper cell types, e.g. between Th1 and Th2. Antagonists of
CC 63954 can be used to modulate immune responses in abnormal situations,
CC e.g. autoimmune disorders, including rheumatoid arthritis, systemic
CC lupus erythematosus (SLE), Hashimoto's autoimmune thyroiditis, as well
CC as acute and chronic inflammatory responses in which T cell activation,
CC expansion, and/or immunological T cell memory play an important role,
CC such as chronic inflammation or tissue rejection. The products can also
CC be used in the treatment of conditions associated with abnormal
CC physiology or development, including abnormal proliferation, e.g.
CC cancerous conditions, or degenerative conditions. The products can also
CC be used for detection, diagnosis and drug screening.
SQ Sequence 285 AA;
```

```
Query Match 21.6%; Score 229.5; DB 1; Length 285;
Best Local Similarity 30.9%; Pred. No. 2e-16;
Matches 67; Conservative 41; Mismatches 86; Indels 23; Gaps 7;

QY 7 QTEQLSRREYSRLQGTGSPQNGEGYPMQSLPEQS--SDALEME-----NGESR 56
D 73 QGDLASLRAELQGHHAELPA--GAGAPKAGLEBPAPVATGAKLFEPAPPEGNSONS 130
QY 57 KRAVLTOKOKKSHVHLVPIINAT--SKDDSDVTEVMQPALRGRLQAGYGVRIODA 115
D 131 NKRAVQGEETVTDCCQLINDSETPTIQKSYTFVPMULSFKGSALAEKENILVYKET 190
QY 116 GYVLLYSQVLEFQDVTFTMGQVVSRE-----GQGRQETLFRICISMP-SHPDRAYNSQYS 168
D 248 AGIAKLEGGDELQALAIPIRENNQISLDGDVTFFGALKL 284
```

```
D 191 GYFPTIGQVLYTDKTYAMGHILQKRVHVEFGDELSVTLFRICIONMPEPLPN---NSCYS 247
QY 169 AGVPHLHQGDLISVTPRRAKLNLSPHGTFGLGVKL 205
D 248 AGIAKLEGGDELQALAIPIRENNQISLDGDVTFFGALKL 284

RESULT 5
W73043 standard: Protein: 285 AA.
ID W73043:
AC W73043:
DE 07-JAN-1999 (first entry)
KW Tumour necrosis factor homologue TL5 protein.
KW Tumour necrosis factor homologue TL5; vaccine; chronic;
KW acute inflammation; arthritis; septicemia; autoimmune disease;
KW inflammatory bowel disease; psoriasis; transplant rejection;
KW graft vs. host disease; infection; stroke; ischaemia;
KW acute respirator disease syndrome; restenosis; brain injury; AIDS;
KW bone disease; cancer; lymphoproliferative disorder; atherosclerosis;
KW Alzheimer's disease.
OS Homo sapiens.
PN EP-869180-A1.
PD 07-OCT-1998.
PF 01-APR-1998: 302526.
PR 03-DEC-1997: US-984396.
PR 02-APR-1997: US-041797.
PA (SMIK ) SMITHKLINE BEECHAM CORP.
PI Hurle MR. Young PR.
DR WPI: 98-508494/44.
DR N-PSDB: V58894.
PT New tumour necrosis factor homologue, TL5 - useful for diagnosis and
PT treatment of Alzheimer's disease, AIDS and cancer
PS Claim 10; Page 18; 23pp; English.
CC The present sequence encodes a tumour necrosis factor homologue TL5
CC polypeptide sequence. TL5 polypeptides and antibodies are useful for
CC identifying compounds which agonise and antagonise TL5, and these can be
CC administered for treatment to inhibit TL5 activity (antagonist) or
CC enhance TL5 activity (agonist). Gene therapy using the expression system
CC can also be used to enhance TL5 activity. Diseases or susceptibility to a
CC disease can be diagnosed by determining the presence or absence of a
CC mutation in the TL5 protein. TL5 polynucleotides are useful for locating
CC genes associated with disease by hybridisation to chromosomes. TL5
CC polypeptides and polynucleotides can be used, especially to raise an
CC immune response (i.e. as vaccines) for the treatment of chronic and acute
CC inflammation, arthritis, septicemia, autoimmune diseases (e.g.
CC inflammatory bowel disease, psoriasis), transplant rejection,
CC graft vs. host disease, infection, stroke, ischaemia, acute respiratory
CC disease syndrome, restenosis, brain injury, AIDS, bone diseases, cancer
CC (e.g. lymphoproliferative disorders), atherosclerosis, and Alzheimers
CC disease.
SQ Sequence 285 AA;
```

```
Query Match 21.6%; Score 229.5; DB 1; Length 285;
Best Local Similarity 30.9%; Pred. No. 2e-16;
Matches 67; Conservative 41; Mismatches 86; Indels 23; Gaps 7;

QY 7 QTEQLSRREYSRLQGTGSPQNGEGYPMQSLPEQS--SDALEME-----NGESR 56
D 73 QGDLASLRAELQGHHAELPA--GAGAPKAGLEBPAPVATGAKLFEPAPPEGNSONS 130
QY 57 KRAVLTOKOKKSHVHLVPIINAT--SKDDSDVTEVMQPALRGRLQAGYGVRIODA 115
D 131 NKRAVQGEETVTDCCQLINDSETPTIQKSYTFVPMULSFKGSALAEKENILVYKET 190
QY 116 GYVLLYSQVLEFQDVTFTMGQVVSRE-----GQGRQETLFRICISMP-SHPDRAYNSQYS 168
D 191 GYFPTIGQVLYTDKTYAMGHILQKRVHVEFGDELSVTLFRICIONMPEPLPN---NSCYS 247
QY 169 AGVPHLHQGDLISVTPRRAKLNLSPHGTFGLGVKL 205
D 248 AGIAKLEGGDELQALAIPIRENNQISLDGDVTFFGALKL 284
```

RESULT 6
W62460 standard; Protein: 285 AA.
ID W62460;
AC W62460;
DT 05-OCT-1998 (first entry)
DE Human T cell surface antigen 63954 protein sequence #1.
KW Human; 63954; primate; rodent; mouse; T cell surface antigen; mammal;
diagnosis; antigen-specific proliferation; cytokine production;
immune response; autoimmune disorder; rheumatoid arthritis;
systemic lupus erythematosus; Hashimoto's autoimmune thyroiditis.
OS Homo sapiens.
PN WO9827114-A2.
PD 25-JUN-1998.
PF 16-DEC-1997; U23321.
PR 17-DEC-1996; US-033601.
PA (SCHE) SCHERING CORP.
PI Gorman DM;
DR MPI: 98-362719/31.
DR N-PSDB: V39984.
PT New isolated polypeptide, 63954 - used to develop products for
treating e.g. autoimmune disorders, inflammation, tissue rejection,
cancer or degenerative conditions
PS Claim 1; Page 57-58; 69pp; English.
CC The present sequence is a human T cell surface antigen, designated
63954. The novel protein designated 63954 is expressed on T cells.
CC Protein 63954 can modulate antigen-specific proliferation and cytokine
production on effector cells and may potentiate immune cell expansion or
apoptosis. 63954 agonists or antagonists may also act as a co-stimulatory
molecule for regulation of T cell mediated cell activation, and may cause
a shift of T helper cell types, e.g. between Th1 and Th2. Antagonists of
63954 can be used to modulate immune responses in abnormal situations,
e.g. autoimmune disorders, including rheumatoid arthritis, systemic
lupus erythematosus (SLE), Hashimoto's autoimmune thyroiditis, as well
as acute and chronic inflammatory responses in which T cell activation,
expansion, and/or immunological T cell memory play an important role,
such as chronic inflammation or tissue rejection. The products can also
be used in the treatment of conditions associated with abnormal
physiology or development, including abnormal proliferation, e.g.
cancerous conditions, or degenerative conditions. The products can also
be used for detection, diagnosis and drug screening.
SQ Sequence 285 AA;

Query Match 21.4%; Score 227.5; DB 1; Length 285;
Best Local Similarity 31.2%; Pred. No. 3.2e-16;
Matches 67; Conservative 40; Mismatches 85; Indels 23; Gaps 7;

QY 9 ELQSLREVSRLQGTGSPQNGEGYPMQSLPEQS--SDALEAME-----NGESRRK 58
DB 75 DLASLRAELQGHHAERKLP--GAGAPRAGLEAPAYTAGKITFEPPAPRGESNSQNSRKN 132
QY 59 RAVLTQKQKQKQSHVHLVPIINAT--SKDDSDVTEVMQPALRGGLQAGQGVRIQDAGV 117
DB 133 RAVQGGEEFTVTOCQDLADSETPTIQKSYTFVPMILSKRGSALEKKNKILVETG 192
QY 118 YLXYSQVLFQDVTFTMGQVSR-----GQGRQETLFRICRSM--SHPDRAVNSQYSAG 170
DB 193 FFLYGGVLYTKTYAMGHLQQRKRVHFGDELSLVTLFRICIQNMPELTPN--NSCYSAG 249
QY 171 VPHLHQGDILSVIIPRAKAKLNSPHGFTLGFVKL 205
DB 250 IAKLEGGDELQALIPRENAQISLDGQDVTFEGALKL 284

RESULT 7
W62462 standard; Protein: 266 AA.
ID W62462;
AC W62462;
DT 05-OCT-1998 (first entry)
DE Human T cell surface antigen 63954 protein sequence #3.
KW Human; 63954; primate; rodent; mouse; T cell surface antigen; mammal;

KW diagnosis; antigen-specific proliferation; cytokine production;
immune response; autoimmune disorder; rheumatoid arthritis;
systemic lupus erythematosus; Hashimoto's autoimmune thyroiditis.
OS Homo sapiens.
PN WO9827114-A2.
PD 25-JUN-1998.
PF 16-DEC-1997; U23321.
PR 17-DEC-1996; US-033601.
PA (SCHE) SCHERING CORP.
PI Gorman DM;
DR MPI: 98-362719/31.
DR N-PSDB: V39986.
PT New isolated polypeptide, 63954 - used to develop products for
treating e.g. autoimmune disorders, inflammation, tissue rejection,
cancer or degenerative conditions
PS Claim 1; Page 63-64; 69pp; English.
CC The present sequence is a human T cell surface antigen, designated
63954. The novel protein designated 63954 is expressed on T cells.
CC Protein 63954 can modulate antigen-specific proliferation and cytokine
production on effector cells and may potentiate immune cell expansion or
apoptosis. 63954 agonists or antagonists may also act as a co-stimulatory
molecule for regulation of T cell mediated cell activation, and may cause
a shift of T helper cell types, e.g. between Th1 and Th2. Antagonists of
63954 can be used to modulate immune responses in abnormal situations,
e.g. autoimmune disorders, including rheumatoid arthritis, systemic
lupus erythematosus (SLE), Hashimoto's autoimmune thyroiditis, as well
as acute and chronic inflammatory responses in which T cell activation,
expansion, and/or immunological T cell memory play an important role,
such as chronic inflammation or tissue rejection. The products can also
be used in the treatment of conditions associated with abnormal
physiology or development, including abnormal proliferation, e.g.
cancerous conditions, or degenerative conditions. The products can also
be used for detection, diagnosis and drug screening.
SQ Sequence 266 AA;

Query Match 20.5%; Score 217; DB 1; Length 266;
Best Local Similarity 29.6%; Pred. No. 3.7e-15;
Matches 64; Conservative 38; Mismatches 74; Indels 40; Gaps 7;

QY 7 QTELRLREVSRLQGTGSPQNGEGYPMQSLPEQS--SDALEAME-----NGESRRK 56
DB 73 QGDLSLRAELQGHHAERKLP--GAGAPRAGLEAPAYTAGKITFEPPAPRGESNSQNSRKN 130
QY 57 KRAVLTQKQKQKQSHVHLVPIINATSKDDSDVTEVMQPALRGGLQAGQGVRIQDAG 116
DB 131 NKRAV-----GQPEETGSYTFVPMILSKRGSALEKKNKILVETG 172
QY 117 YLXYSQVLFQDVTFTMGQVSR-----GQGRQETLFRICRSM--SHPDRAVNSQYSAG 169
DB 173 YFFLYGGVLYTKTYAMGHLQQRKRVHFGDELSLVTLFRICIQNMPELTPN--NSCYSAG 229
QY 170 GVPHLHQGDILSVIIPRAKAKLNSPHGFTLGFVKL 205
DB 230 GAKLEGGDELQALIPRENAQISLDGQDVTFEGALKL 265

RESULT 8
W62463 standard; Protein: 150 AA.
ID W62463;
AC W62463;
DT 05-OCT-1998 (first entry)
DE Mouse T cell surface antigen 63954 protein sequence.
KW Human; 63954; primate; rodent; mouse; T cell surface antigen; mammal;
diagnosis; antigen-specific proliferation; cytokine production;
immune response; autoimmune disorder; rheumatoid arthritis;
systemic lupus erythematosus; Hashimoto's autoimmune thyroiditis.
OS Mus sp.
PN WO9827114-A2.
PD 25-JUN-1998.
PF 16-DEC-1997; U23321.
PR 17-DEC-1996; US-033601.
PA (SCHE) SCHERING CORP.

PI Gorman DM;
DR WPI: 98-362719/31.
DR N-PSDB: V39987.
PT New isolated polypeptide, 63954 - used to develop products for
PT treating e.g. autoimmune disorders, inflammation, tissue rejection,
PT cancer or degenerative conditions
PS Claim 1; Page 66; 69pp; English.
CC The present sequence is a mouse T cell surface antigen, designated
CC 63954. The novel protein designated 63954 is expressed on T cells.
CC Protein 63954 can modulate antigen-specific proliferation and cytokine
CC production on effector cells and may potentiate immune cell expansion or
CC apoptosis. 63954 agonists or antagonists may also act as a co-stimulatory
CC molecule for regulation of T cell mediated cell activation, and may cause
CC a shift of T helper cell types, e.g. between Th1 and Th2. Antagonists of
CC 63954 can be used to modulate immune responses in abnormal situations,
CC e.g. autoimmune disorders, including rheumatoid arthritis, systemic
CC lupus erythematosus (SLE), Hashimoto's autoimmune thyroiditis, as well
CC as acute and chronic inflammatory responses in which T cell activation,
CC expansion, and/or immunological T cell memory play an important role,
CC such as chronic inflammation or tissue rejection. The products can also
CC be used in the treatment of conditions associated with abnormal
CC physiology or development, including abnormal proliferation, e.g.
CC cancerous conditions, or degenerative conditions. The products can also
CC be used for detection, diagnosis and drug screening.
SQ Sequence 150 AA;

Query Match 18.6%; Score 197; DB 1; Length 150;
Best Local Similarity 36.3%; Pred. No. 2.2e-13;
Matches 45; Conservative 26; Mismatches 43; Indels 10; Gaps 3;

OY 89 TEVMMQPALRRRGQIAQGYGVRIQAGVYLLYSQVLFQDVFYTMQGVYSRE-----GQ 142
DB 29 TFFVWLISKRNALAEKKNKTVVROTGFYISQVLYDPLFAMCHVQIRKKVHVGDE 88
OY 143 GHOETLFRCIRSMPSHPDRAYNSQYACVPHOGDILSVIIPRAKILNLSPHGTFPG 201
DB 89 LSLVTLFRICIQMFKTLPR---NSCYSAGIARLEEDDEIQLAIPRENAQISRNQDITFG 145
OY 202 FVKL 205
DB 146 ALKL 149

RESULT 9
R44145
ID R44145 standard; Protein: 155 AA.
AC R44145;
DT 18-MAY-1994 (first entry)
DE TNF RGD muteln #1.
KW Wild type; tumour necrosis factor; RGD motif; human; TNF; muteln;
KW antitumour activity; cancer; metastasis.
OS Homo sapiens.
FH Key Location/Qualifiers
FT misc_difference 56 /note= "Encoded by ATG"
FT J05271289-A.
PD 19-OCT-1993.
PE 17-MAR-1992; 108342.
PR 17-MAR-1992; JP-108342.
RA (ISHH) ISHIMURA SANGYO KAISHA LTD.
DR WPI: 93-365243/46.
PT Tumour necrosis factor muteln sequence - and recombinant plasmid
PT contg. DNA coding for it, and transformed microorganisms used for
PT its prodn.
PS Claim 2; Page 37; 56pp; English.
CC The sequences given in R44145-52 represents tumour necrosis factor
CC (TNF) muteins in which amino acids 1-8 of wild type TNF are substituted
CC by an amino acid sequence comprising 3-16 amino acids and containing
CC at least one Arg-Gly-Asp motif. These human TNF muteins have the same
CC antitumour activity as wild type TNF and show no cancer metastasis
CC promoting activity. These muteins can be produced by recombinant DNA
CC techniques and may be used in drug compositions.

SQ Sequence 155 AA;

Query Match 9.4%; Score 99.5; DB 1; Length 155;
Best Local Similarity 30.1%; Pred. No. 0.0039;
Matches 41; Conservative 18; Mismatches 48; Indels 29; Gaps 7;

OY 80 ARSKDSD--VTEVMMQPAL-----RRGRGIAQGYGVRIQAGVYLLYSQV 124
DB 1 SSRGDSKPVAKHVVANPAEGOLQMLNRRANLLANGVELRDNOLVPSSEGILYISQV 60
OY 125 LFQD-----VFYTMQGV-VSRGQGRQETLFR--CIRSMPSHPDRA--YNSQYSAG 170
DB 61 LKRGQCSPTHVLTHTMRINAVSYOTKYNLSAIRSPQRETPDEGAEPKWPETPLDG 120
OY 171 VPHLHOGDILSVIIPR 186
DB 121 VFQLEKGRDLSAQINR 136

RESULT 10
P60656
ID P60656 standard; Protein: 158 AA.
AC P60656;
DT 28-JUL-1991 (first entry)
DE Sequence of mature human tumour necrosis factor (hTNF) muteln
DE Ser 69 in PAW/31.
KW Antitumour; anticancer.
OS Homo sapiens.
PN WO8604606-A.
PD 14-AUG-1986.
PE 03-FEB-1986; U00236.
PR 19-OCT-1982; US-435154.
PR 15-APR-1983; US-486162.
PR 20-DEC-1983; US-564224.
PR 15-OCT-1984; US-661026.
PR 07-FEB-1985; US-698939.
PR 07-FEB-1985; US-695934.
PA (CETU) CETUS CORP.
PI Mark DF, Lin LS, Lu SDY, Wang AM;
DR WPI: 86-225458/34.
DR N-PSDB: N60558.
PT New synthetic muteins of human tumour necrosis factor protein -
PT are Obtd. by direct mutagenesis and retain antitumour activity
PS Disclosure; Fig 3a; 47pp; English.
CC The sequence encoding TNF produced by the promyelocytic leukemia
CC cell line (HL-60, ATCC no CCL240) has been cloned and expressed in
CC E.coli (see N60557). Neither of the cysteine residues (69 and 101) in
CC the TNF sequence appears to be involved in disulphide linkages. The
CC CC patentors claim a novel synthetic muteln of a biologically active
CC hTNF protein, having at least one cysteine residue free from a
CC disulphide link and non-essential to the activity and having at
CC least one of the cysteine residues deleted or replaced by another AA.
CC Plasmid pAM731 (Ser 69) is claimed.
SQ Sequence 158 AA;

Query Match 9.3%; Score 98.5; DB 1; Length 158;
Best Local Similarity 28.1%; Pred. No. 0.0031;
Matches 39; Conservative 20; Mismatches 53; Indels 27; Gaps 6;

OY 75 LVPINATSKDSDVTEVMMQPAL-----RRGRGIAQGYGVRIQAGVYLLYSQV 121
DB 1 MVRSSSRTPSDKPVAKHVVANPAEGOLQMLNRRANLLANGVELRDNOLVPSSEGILYI 60
OY 122 SQVLFQD-----VFYTMQGV-VSRGQGRQETLFR--CIRSMPSHPDRA--YNSQI 167
DB 61 SQVLFQDQSPSTHVLTHTMRINAVSYOTKYNLSAIRSPQRETPDEGAEPKWPETPL 120
OY 168 SAGVPHLHOGDILSVIIPR 186
DB 121 LGGVFOLEKGRDLSAQINR 139

This Page Blank (uspto)

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 13, 1999, 20:59:04 ; Search time 33.52 Seconds

(Without alignments)
229.098 Million cell updates/sec

Title: US-09-212-270-5

Perfect score: 1061
Sequence: 1 MWOLNQQTFLQSLRRVSRLL.....RARAKLNSPHGTFLEGVKL 205

Scoring table: BLOSUM62

Searched: 116738 seqs, 37460341 residues

Database: PIR_58.*

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|----------------------|
| 1 | 96 | 9.0 | 235 | 2 | JU0029 | tumor necrosis fac |
| 2 | 92 | 8.7 | 233 | 2 | S22052 | tumor necrosis fac |
| 3 | 91.5 | 8.6 | 233 | 1 | OMHUN | tumor necrosis fac |
| 4 | 88 | 8.3 | 185 | 2 | S52715 | tumor necrosis fac |
| 5 | 87.5 | 8.2 | 235 | 1 | QWMSN | tumor necrosis fac |
| 6 | 87 | 8.2 | 233 | 2 | S24642 | tumor necrosis fac |
| 7 | 86.5 | 8.2 | 233 | 2 | S11688 | tumor necrosis fac |
| 8 | 86 | 8.1 | 197 | 2 | JH0309 | tumor necrosis fac |
| 9 | 85.5 | 8.1 | 205 | 1 | OMHUX | lymphotoxin alpha |
| 10 | 85.5 | 8.1 | 234 | 2 | Q01344 | tumor necrosis fac |
| 11 | 85.5 | 8.1 | 234 | 2 | A25451 | tumor necrosis fac |
| 12 | 82.5 | 7.8 | 235 | 2 | S15490 | tumor necrosis fac |
| 13 | 81 | 7.6 | 392 | 2 | I54491 | cell surface antig |
| 14 | 80.5 | 7.6 | 357 | 2 | D71262 | probable hemolysin |
| 15 | 80 | 7.5 | 234 | 2 | JH0529 | tumor necrosis fac |
| 16 | 80 | 7.5 | 233 | 2 | S13114 | tumor necrosis fac |
| 17 | 80 | 7.5 | 707 | 2 | S78538 | site-specific reco |
| 18 | 78 | 7.4 | 232 | 2 | S12606 | tumor necrosis fac |
| 19 | 78 | 7.4 | 281 | 2 | I38707 | Fas ligand - human |
| 20 | 77 | 7.3 | 505 | 2 | I37206 | protein-tyrosine k |
| 21 | 77 | 7.3 | 193 | 2 | S06192 | tumor necrosis fac |
| 22 | 76 | 7.2 | 171 | 2 | B44126 | ribosomal protein |
| 23 | 75.5 | 7.1 | 484 | 2 | S48403 | sc/svp protein - yea |
| 24 | 75.5 | 7.1 | 753 | 2 | S18689 | FKH1 protein - yea |
| 25 | 74 | 7.0 | 393 | 2 | E64485 | phosphoribosylglyc |
| 26 | 74 | 7.0 | 1068 | 2 | S60855 | pleiotropic drug r |
| 27 | 74 | 7.0 | 486 | 2 | S00855 | hypothetical prote |
| 28 | 73.5 | 6.9 | 202 | 2 | B27303 | tumor necrosis fac |
| 29 | 73 | 6.9 | 670 | 2 | S77387 | nuclear transport |
| 30 | 73 | 6.9 | 1996 | 2 | F71405 | probable TMV resis |
| 31 | 72 | 6.8 | 726 | 1 | FC6430 | translation elonga |
| 32 | 72 | 6.8 | 360 | 1 | FC6430 | translation elonga |
| 33 | 72 | 6.8 | 1586 | 2 | S39580 | HRM protein - hum |
| 34 | 72 | 6.8 | 1572 | 2 | S45251 | SNF2alpha protein |
| 35 | 72 | 6.8 | 749 | 2 | S64780 | hypothetical prote |
| 36 | 72 | 6.8 | 1120 | 2 | S67208 | hypothetical prote |
| 37 | 71.5 | 6.7 | 513 | 1 | RGECAV | transcription regu |
| 38 | 71.5 | 6.7 | 775 | 2 | B70449 | phenylalanyl-tRNA |
| 39 | 71.5 | 6.7 | 337 | 2 | S21405 | porin - Neisseria |

| | | | | | | |
|----|------|-----|-----|---|--------|--------------------|
| 40 | 71.5 | 6.7 | 608 | 2 | A64992 | sensor protein Ato |
| 41 | 71 | 6.7 | 416 | 2 | JN0006 | nerve growth facto |
| 42 | 71 | 6.7 | 717 | 2 | S31035 | retrovirus-related |
| 43 | 71 | 6.7 | 717 | 2 | S31034 | retrovirus-related |
| 44 | 71 | 6.7 | 179 | 2 | B64861 | hypothetical prote |
| 45 | 71 | 6.7 | 638 | 2 | S67605 | hypothetical prote |

ALIGNMENTS

RESULT 1

tumor necrosis factor alpha precursor - rat
N/Alternate names: cachectin, TNF alpha
C/Species: Rattus norvegicus (Norway rat)
C/Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 20-Mar-1998
R/Shirai, T.; Shimizu, N.; Horiguchi, S.; Ito, H.
Agric. Biol. Chem. 53, 1733-1736, 1989
A/Title: Cloning and expression in Escherichia coli of the gene for rat tumor necrosi
A/Reference number: JU0029
A/Accession: JU0029
A/Molecule type: DNA
A/Residues: 1-235 <SHI>
R/Kwon, J.; Chung, I.Y.; Benveniste, E.N.
Gene 132, 227-236, 1993
A/Title: Cloning and sequence analysis of the rat tumor necrosis factor-encoding gene
A/Reference number: JN0868
A/Accession: JN0868
A/Molecule type: DNA
A/Residues: 1-235 <KWO>
A/Cross-references: GB:L00981; NID:q205253; PID:q205254
R/Estler, H.C.; Grewe, M.; Gausling, R.; Pavlovic, M.; Decker, K.
Biol. Chem. Hoppe-Sevler 373, 271-281, 1992
A/Title: Rat tumor necrosis factor-alpha. Transcription in rat Kupffer cells and in v
A/Reference number: S21674; M01D:92329007
A/Accession: S21674
A/Molecule type: mRNA
A/Residues: 1-38, 'P', '40-162, 'T', '164-201, 'S', '203-235 <EST>
A/Cross-references: GB:X6539; GB:S40199; NID:q395369; PID:q395370
C/Comment: Tumor necrosis factor is secreted by macrophages in response to endotoxin
C/Genetics:
A/Gene: TNF-alpha
A/Intons: 62/3; 81/1; 97/1
C/Superfamily: tumor necrosis factor
C/Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane prote
F;80-225/Product: tumor necrosis factor #status predicted <MAT>
F;86/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.0%; Score 96; DB 2; Length 235;
Best Local Similarity 27.5%; Pred. No. 0.052;
Matches 58; Conservative 29; Mismatches 68; Indels 56; Gaps 15;

| | | | |
|----|-----|--|---------------------------|
| QY | 22 | GTGPGSONGEGYPMQSLPEQSSDALLEAMENGERSRKRRAVLTOKRKHOS--- | VLHLP 77 |
| DB | 54 | GVIGPKNE-EKFP-NGPLISSMA----- | QTLTLRSSSSONSNDKRAHVVA 97 |
| QY | 78 | INATSKDSVTEVMOPALRRGRLOAGYGVRA---- | IQDAGVLLYLYSOVLFQD--- 129 |
| DB | 98 | NHQAE-----QLEW--LSQANALLANGMDLKNOLVVPADLILYSQVLEFGGCGP | 149 |
| QY | 129 | --VFETMGVYVREGQGRQ--TLFRGIRS--MPSHEDRA----- | YNSQVSAGVPPLHOG 177 |
| DB | 150 | DYVLLT--HTVSRPAISYQKRVSLSAIKSPCKRDTPEGALKEWPEPMYGVGVLEKG | 207 |
| QY | 178 | DLIS--VILPRARAKLNSPHG--TFLEGVKL 205 | |
| DB | 208 | DLISAEVNLKP---YLDITSGGVYFVIAL 235 | |

RESULT 2

A:Title: Haploctypic polymorphisms of the TNFb gene.
A:Reference number: 154482; MUID:91139175
A:Accession: 154482
A:Status: translation not shown; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-124, 'P', 126-205 <RES>
A:Cross-references: GB:M55913; NID:g339742; PID:g339743
A:Experimental source: ancestral haplotype 57.1
A>Note: 59-Asn was also found (ancestral haplotype 8.1)
R:Giray, P.W.; Aggarwal, B.B.; Benton, C.V.; Bringman, T.S.; Henzel, W.J.; Jarrett, J.A.,
Nature 312, 721-724, 1984
A:Title: Cloning and expression of cDNA for human lymphotoxin, a lymphokine with tumour
A:Reference number: A33350; MUID:85086243
A:Accession: A93350
A:Molecule type: mRNA
A:Residues: 1-205 <GRA>
A:Cross-references: GB:X01393; NID:g34444; PID:g34445
A:Experimental source: Lymphoblastoid cell line RPMI-1788
R:Coedel, D.V.; Aggarwal, B.B.; Gray, P.W.; Leung, D.W.; Nedvin, G.E.; Palladino, M.A.,
Cold Spring Harb. Symp. Quant. Biol. 51, 597-609, 1986
A:Title: Tumor necrosis factors: gene structure and biological activities.
A:Reference number: A32877; MUID:87217059
A:Accession: B32877
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 35-205 <GOE>
R:Kobayashi, Y.; Miyamoto, D.; Asada, M.; Obinata, M.; Osawa, T.
J. Biochem. 100, 727-733, 1986
A:Title: Cloning and expression of human lymphotoxin mRNA derived from a human T cell hybridoma
A:Reference number: A91906; MUID:87057135
A:Accession: A91906
A:Molecule type: mRNA
A:Residues: 1-59, 'N', 61-205 <KOB>
A:Cross-references: GB:PD0102; NID:g219913; PID:d1000508; PID:g219914
A>Note: the authors translated the codon TAT for residue 156 as Thr and ACC for residue
R:Fukuda, S.; Ando, S.; Sanou, O.; Tanai, M.; Fujii, M.; Masaki, N.; Nakamura, K.I.; Adachi,
Lymphokine Res. 7, 175-185, 1988
A:Title: Simultaneous production of natural human tumor necrosis factor-alpha, -beta and ACC for residue
A:Reference number: A61478
A:Accession: A61478
A:Molecule type: protein
A:Residues: 56-79,86-95, 'X', 97, 'X', 99,119-151, 'XX',154-162, 'X',164, 'X',166, 'X',168, 'X',1
R:Voigt, C.G.; Maurer-Fogy, I.; Adolif, G.R.
FEBS Lett. 314, 85-88, 1992
A:Title: Natural human tumor necrosis factor beta (lymphotoxin). Variable O-glycosylation
A:Reference number: S26951
A:Accession: S26951
A:Molecule type: protein
A:Residues: 35-59, 'N', 61-205 <VOI>
A>Note: 60-Thr was also found
R:Fukushima, K.; Watanabe, H.; Takeo, K.; Nomura, M.; Asahi, T.; Yamashita, K.
Arch. Biochem. Biophys. 304, 144-153, 1993
A:Title: N-linked sugar chain structure of recombinant human lymphotoxin produced by CHO
A:Reference number: S34742
A:Accession: S34742
A:Contents: annotation
A:Comment: secreted from mitogen-activated lymphocytes within 1-2 days after induction,
while having no detrimental effect on normal cells. It can also act synergistically with
A:Comment: This protein and TNF-alpha (tumor necrosis factor) are the products of differ
A:Comment: local activities but are produced by different cell types and have different induction k
A:Gene: GDB:LTA; LT; TNFB
A:Cross-references: GDB:120442; OMIM:153440
A:Map position: 6p21.3-6p21.3
A:Introns: 33/3; 69/1
A>Note: The first intron occurs before the initiator codon
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; homotrimer; lymphokine; macrophage
F:1-4/Domain: signal sequence #status predicted <SIG>
F:35-205/Product: lymphotoxin #status predicted <MAP>
F:41/Binding site: carbohydrate (Asn) (covalent) #status experimental
F:96/Binding site: carbohydrate (Asn) (covalent) #status experimental

| | Query Match | 8.1% | Score 85.5: | DB 1: | Length 205: |
|----|---|-------|--------------|----------|-----------------------------------|
| | Best Local Similarity | 23.9% | Pred. | No.0.44: | |
| | Matches | 42: | Conservative | 19: | Mismatches 50: Indels 65: Gaps 9: |
| Oy | KOHSHVHLVPINATSK-----DDSDTEVMWOPALRRG---RGDAQGYGVRIODAGV | 117 | :::: | : :: | : |
| Dd | RQHPKMHLL--HSTLKPRAAHLIDPSKSOSLIMRANTDRAFLQDFSLSNNSLLPTSGI | 106 | :: | : :: | : |
| Oy | YLXYSQVLFDVYTTMGOVNRSREGOROEELFRCLRSMP---SHPDRAYNSOYSAGVP-- | 173 | | : :: | : |
| Dd | VFYVSQVFES-----GRAVSPKA-----TSBPLYLAHEVOLFSQQYPHPVPLL | 149 | : :: | : :: | : |
| Oy | 173-----HLHGDIISLV---IPRARAKILNSPHGTFFLG | 201 | | : :: | : |
| Dd | SQAQNVYEGLOEPWLMSNYHGAAFQLQTGDGLSTRTDGIPH---LVLSPSPIVEFG | 201 | | : :: | : |

```

RESULT 10
J01344
tumor necrosis factor alpha precursor - horse
N:Alternate names: cachectin; TNF alpha
C:Species: Equus caballus (domestic horse)
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 08-Sep-1997
C:Accession: J01344
R:Su, X.; Morris, D.D.; McGraw, R.A.
Gene 107, 319-321, 1991
A:Title: Cloning and characterization of gene TNF alpha encoding equine tumor necrosis
A:Reference number: J01344; MUID:92084125
A:Accession: J01344
A:Molecule type: DNA
A:Residues: 1-234 <SU>
A:Cross-references: GB:M4087; NID:g164244; PID:g164245
C:Comment: This protein is an important proximal mediator of endotoxemia.
C:Genetics:
A:Gene: TNF-alpha
A:Introns: 62/3; 79/1; 95/1
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane prote
F:78-234/Product: tumor necrosis factor alpha #status predicted <TV>

Query Match 8.1%; Score 85.5; DB 2; Length 234;
Best Local Similarity 27.3%; Pred. No. 0.52;
Matches 41; Conservative 14; Mismatches 38; Indels 57; Gaps 8;

OY 78 INATSKDSD--YEVMMQPALR-----RGRGLAQSGYVRID-----AGVYLLYS 122
      ::::: || |::: | ||::: | ||::: |
DB 78 LRSRRTPSDKPVAAHVAVNPQAEQQLMLSGRANLANGVKLDYDNLVPLDGILYIS 137
      ||| |::: |::: |::: |::: |::: |
OY 123 QVLFQDYVFTFGQVVYSRGQGRQET-----LFRGIRSMPSH---P 159
      ||| |::: |::: |::: |::: |::: |
DB 138 QVLF-----KGQCPSTHVLTLTISRLANYSKYNLSAITS-FCHTESP 183
      ||| |::: |::: |::: |::: |::: |
OY 160 DRA-----YNSQYSAGVPHLHQGDILSYII 184
      ::| | | | | | | | | | |
DB 184 EQAEKAPWYEPYILGVGFQLEKGDQLSAEI 213
      ::| | | | | | | | | | |

RESULT 11
A25451
tumor necrosis factor alpha precursor - rabbit
N:Alternate names: cachectin; TNF alpha
C:Species: Oryctolagus cuniculus (domestic rabbit)
C:Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 20-Mar-1998
C:Accession: A25454; A23551; J50727
R:Itto, H.; Yamamoto, S.; Kuwoda, S.; Sakamoto, H.; Kajihara, J.; Kiyota, T.; Hayashi,
DNA 5, 149-156, 1986
A:Title: Molecular cloning and expression in Escherichia coli of the cDNA coding for
A:Reference number: A25454; MUID:86219711
A:Accession: A25454
A:Molecule type: mRNA
A:Residues: 1-234 <TT>
A:Cross-references: GB:M12845; NID:g165759; PID:g165760

```

R:Rico, H.; Shirai, T.; Yamamoto, S.; Akira, M.; Kawahara, S.; Todd, C.W.; Wallace, R.B.
DNA 5, 157-165, 1986
A:Title: Molecular cloning of the gene encoding rabbit tumor necrosis factor.
A:Reference number: A25451; MUID:86219712
A:Accession: A25451
A:Molecule type: DNA
A:Residues: 1-234 <A12>
A:Note: this sequence differs from that shown in having a Gln inserted between residues
R:Shakhov, A.N.; Kuprash, D.V.; Azizov, M.M.; Jongeneel, C.V.; Nedospasov, S.A.
Gene 95, 215-221, 1990
A:Title: Structural analysis of the rabbit TNF locus, containing the genes encoding TNF
A:Reference number: JH0309; MUID:91065534
A:Accession: J50727
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-62,'O',63-234 <SHA>
A:Cross-references: GB:M60340; GB:M35326; NID:g165754; PID:g165756
C:Genetics:
A:Introns: 62/3; 80/1; 96/1
C:Superfamily: tumor necrosis factor
C:Keywords: cytokine; cytotoxin; glycoprotein; lymphokine; macrophage; membrane protein
F:1-81/Domain: signal sequence #status predicted <Sig>
F:82-234/Product: tumor necrosis factor #status predicted <Mat>

| | | | | |
|--------------------------|--------|-----------------|------------|-------------|
| Query Match | 8.18; | Score 85.5; | DB 2; | Length 234; |
| Best Local Similarity | 23.98; | Pred. NO. 0.52; | | |
| Matches 47; Conservative | 28; | Mismatches 53; | Indels 69; | Gaps 12; |

| | | | | | | | |
|----|-----|--|---|---------------------------------|------------------|-------------|-----|
| QY | 61 | VLTQKOKROHSHVL | -----VINTSK | - | DDSDYTEVMQNAL | -----RRRGGL | 103 |
| | | : : : : | : : : : | | : : : : | : : : : | |
| Db | 55 | VIGQOEESPNNHLVNPVAMQTLTRLSRALSDKPLAHVAVANQVGCQLQWLMSQRANAL | | | | | 114 |
| QY | 104 | QAQGVGVRIOD | -----AGVYLLYSQVLEFQDYVFTMGQVVSREGQ | ----- | | | 144 |
| | | : : : : | : : : : | | : : : : | : : : : | |
| Db | 115 | LAN--GKMLTDNQLVPPADGIXLYLSQVLF | | | SGGCRSRVYLLTHYSR | | 159 |
| QY | 144 | -----ROETLFCIRSMPSH | -----PDRA | -----YNSQYSAQVPHLHQDILLVYIIPARA | | | 189 |
| | | : : : : | : : : : | : : : : | : : : : | : : : : | |
| Db | 160 | FAVSPVKNVNLISAKIS | -PCHREMPPEAEAPMAWEPIYLGQVFQLEKGRDLRSTEVNQPE | | | | 218 |
| QY | 190 | KLNLSPHG | -FTLGFVKL | 205 | | | |
| | | : : : : | : : : : | | | | |
| Db | 218 | YLDLAESGVYFGIIAL | 234 | | | | |

RESULT 12
154490
tumor necrosis factor - white-footed mouse
C:Species: Peromyscus leucopus (white-footed mouse)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 23-Feb-1997
C:Accession: 154490
R:Crew, M.D.; Filipowsky, M.E.
Immunogenetics 35, 351-353, 1992
A:Title: Sequence of the tumor necrosis factor/cachectin (TNF) gene from Peromyscus leuc
A:Reference number: 154490; M0ID:92216012
A:Accession: 154490
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-235 <RES>
A:Cross-references: GB:M59233; NID:g202506; PID:g202507
C:Genetics:
A:Gene: lTNF
A:Introns: 62/3; 81/1; 97/1
A:Superfamily: tumor necrosis factor

| | | | | | |
|----|--|-------|--------------|-------|----------------|
| | Query Match | 7.8% | Score 82.5; | DB 2; | Length 235; |
| | Best Local Similarity | 27.9% | Pred. No. 1; | | |
| | Matches | 56; | Conservative | 25; | Mismatches 65; |
| | | | | | Indels 55; |
| | | | | | Gaps 14. |
| OY | 22 GTGGGQQNGEGPMNSLPDQSSDALEAMENGRSKRRRAVLÖKÖKKÖKS---- <td>77</td> <td></td> <td></td> <td></td> | 77 | | | |
| | : : | | : | : | : |

```

Db      54 GVIGP-QREEFP--NNLPIIGSMA-----QTLIRSSQNSSDKPVAAHV-  97

QY      78 INATSKDDSDVETVMQPALRG-RGIAQAGYVR-----IODAGVLLYSQVLFQD---  129
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : :
Db      97 --ANHQW---QLEW---LSRGALLANGMDLKQWLVIPADGLYLVYSQVLFGRGQG  148
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : :

QY      129 -----VFTMQQ-VYSKEGGQROGTLR--CISMSHPHRA--YNSQYSAGVPHLHOGD  178
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : :
Db      149 SSYYLLHTHTYSAFVNSYEDKYNLLSAIKSPCPRETPEGSELKPMYEPYILYGVFOLEKGD  208
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : :

QY      179 ILSVILPRAKINLSPHGTF 199
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : :
Db      209 RLS-----AEVNLPKYLD 222
      - - - - - : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT      13
154491
cell surface antigen - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Feb-1997
C:Accession: 154491
R:Kristits, M.J.; Kunz, H.W.; Cortese Hassett, A.L.; Gill, T.J.
Immunogenetics 35, 365-377, 1992
A:Title: Genomic DNA sequence and organisation of a TL-like gene in the grc-G/C region
A:Reference numbers: 154491; MUID:92250119
A:Accession: 154491
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-392 <RES>
A:Cross-references: GB:M74822; NID:g205461; PID:g205462
C:Genetics:
A:Introns: 31/1; 121/1; 205/1; 297/1; 336/1; 347/1; 362/3
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
:Keywords: surface antigen

```

| Query Match | 7.6% | Score 81 | DB 2 | Length 392 |
|-----------------------|-------|--|------|----------------------------------|
| Best Local Similarity | 22.0% | Pred. NO. 2.8 | | |
| Matches | 37 | Conservative | 22 | Mismatches 61; Indels 48; Gaps 8 |
| QY | 34 | PMOSLPEOSSDLEAMENGERSRKRRAVLTKOKKHSHVLTLPINATSKDDSDTEVVM | 93 | |
| | | | | |
| | | | | |
| Db | 80 | PM-----QGEDELWE-----QYHKKTKDQGSARNIMMLVRYTKNSMDFHT--LQW | 127 | |
| | | | | |
| QY | 94 | QPALRRGGGLAQGYGVRIQAGVYLLYSQVLEFDY-----FTMGQVVS | 138 | |
| | | | | |
| Db | 128 | Q-----QCGDGV--SDGRLLHWYDQLAFDGVDPHTLNKDLRFWTAMTSTVAQISQ | 175 | |
| | | | | |
| QY | 139 | REGCGROETLFCRISMSHPDRATNSQYSAGVPLH-----QGD I 179 | | |
| | | | | |
| Db | 176 | PELEKRLD--NCSELLQKVPKEKERILLRSDPPAHVATHROPREPQGV 221 | | |
| | | | | |

RESULT 14

D71262

Probable hemolysin - syphilis spirochete

C:Species: *Treponema pallidum* subsp. *pallidum* (syphilis spirochete)

C:Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 07-Aug-1998

C:Accession: D71262

R:Fraser, C.M.; Norris, S.J.; Weinstein, G.M.; White, O.; Sutton, G.G.; Dodson, J.; Rison, J.; Khalak, H.; Richardson, D.; Howell, J.R.; Chidambaram, M.; Utterback, T.; Chey, L.; Weidman, J.; Smith, H.O.; Venter, J.C.

A:Title: Complete genome sequence of *Treponema pallidum*, the syphilis spirochete

A:Reference number: A71250

A:Accession: D71262

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1357 <COL>

A:Cross-references: GB:AE001262; GB:AE000520; NID:g3322254; PTD:g3323257

A:Experimental source: strain Nicholas

C:Genetics:

A:Gene: TP0936

This Page Blank (uspio)

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 1999, 14:43:39 ; Search time 25.8 Seconds

(without alignments)
213.266 Million cell updates/sec

Title: US-09-212-270-5

Perfect score: 1061

Sequence: 1 MVOLTOQTQELQSLRREVSRL.....RARAKLNLSPHGTFLGEVYKL 205

Scoring table: BLOSUM62

Searched: 74019 seqs, 26840295 residues

Database : SwissProt_36.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|-------------|
| 1 | 96 | 9.0 | 235 | 1 | TNFA_RAT |
| 2 | 95.5 | 9.0 | 234 | 1 | TNFA_CAVPO |
| 3 | 92 | 8.7 | 233 | 1 | TNFA_PAPSP |
| 4 | 91.5 | 8.6 | 233 | 1 | TNFA_HUMAN |
| 5 | 90.5 | 8.5 | 235 | 1 | TNFA_HUMAN |
| 6 | 88 | 8.3 | 229 | 1 | TNFA_RABIT |
| 7 | 87.5 | 8.2 | 235 | 1 | TNFA_MOUSE |
| 8 | 87 | 8.2 | 233 | 1 | TNFA_MOUSE |
| 9 | 86.5 | 8.2 | 233 | 1 | TNFA_BOVIN |
| 10 | 86 | 8.1 | 233 | 1 | TNFA_FELCA |
| 11 | 86 | 8.1 | 233 | 1 | TNFA_MACMU |
| 12 | 86 | 8.1 | 197 | 1 | TNFA_RABIT |
| 13 | 85.5 | 8.1 | 234 | 1 | TNFA_HORSE |
| 14 | 85.5 | 8.1 | 205 | 1 | TNFA_HUMAN |
| 15 | 83.5 | 7.9 | 233 | 1 | TNFA_CANFA |
| 16 | 82.5 | 7.8 | 235 | 1 | TNFA_PERLE |
| 17 | 80 | 7.5 | 234 | 1 | TNFA_SHEEP |
| 18 | 80 | 7.5 | 233 | 1 | TNFA_TRIYU |
| 19 | 78 | 7.4 | 281 | 1 | FASL_HUMAN |
| 20 | 78 | 7.4 | 232 | 1 | TNFA_PIG |
| 21 | 77.5 | 7.3 | 291 | 1 | TRAI_MOUSE |
| 22 | 77 | 7.3 | 504 | 1 | BLK_HUMAN |
| 23 | 77 | 7.3 | 193 | 1 | TNFA_CAPHI |
| 24 | 76 | 7.2 | 171 | 1 | RS4_HALMA |
| 25 | 75.5 | 7.1 | 484 | 1 | FHL1_YEAST |
| 26 | 74 | 7.0 | 496 | 1 | IBMP_CERV |
| 27 | 73.5 | 6.9 | 1063 | 1 | PDR1_YEAST |
| 28 | 73.5 | 6.9 | 638 | 1 | FTSH_HELPF |
| 29 | 73.5 | 6.9 | 202 | 1 | TNFA_MOUSE |
| 30 | 73 | 6.9 | 670 | 1 | NTRC_SYNY3 |
| 31 | 72 | 6.8 | 726 | 1 | EF2_METUA |
| 32 | 72 | 6.8 | 360 | 1 | RFL_ECOLI |
| 33 | 72 | 6.8 | 1586 | 1 | SN22_HUMAN |
| 34 | 71.5 | 6.7 | 608 | 1 | ATOS_ECOLI |
| 35 | 71.5 | 6.7 | 1863 | 1 | BRC1_HUMAN |
| 36 | 71.5 | 6.7 | 337 | 1 | OMB_NETLA |
| 37 | 71.5 | 6.7 | 513 | 1 | TTRR_ECOLI |
| 38 | 71 | 6.7 | 416 | 1 | NGFR_CHICK |
| 39 | 71 | 6.7 | 175 | 1 | YCFE_ECOLI |
| 40 | 71 | 6.7 | 378 | 1 | ZP47_BRARE |
| 41 | 70.5 | 6.6 | 325 | 1 | IF34_HUMAN |
| 42 | 70.5 | 6.6 | 204 | 1 | TNFB_PIG |
| 43 | 70.5 | 6.6 | 554 | 1 | YJUK_ECOLI |

ALIGNMENTS

44 70 6.6 361 1 COOH_RHQRU
45 70 6.6 278 1 FASL_RAT

P31895 rhodospirill
P36940 rattus norv

RESULT 1

TNFA_RAT
ID TNFA_RAT STANDARD: PRT: 235 AA.
AC P16599;
DT 01-AUG-1990 (REL. 15, CREATED)
DT 01-AUG-1990 (REL. 15, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS RATTUS NORVEGICUS (RAT).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-TESTIS;
RX MEDLINE: 94040766.
RA KWON J.; CHUNG I.Y.; BENVENISTE E.N.;
RL GENE 132:227-236(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA SHIRAI T., SHIMIZU N., HORIGUCHI S., ITO H.;
RL AGRIC. BIOL. CHEM. 53:1733-1736(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92329007.
RA ESTLER H.C., GRENE M., GAUSSLING R., PAVLOVIC M., DECKER K.;
RL BIOL. CHEM. HOPPE-SEYLER 373:271-281(1992).
RN [4]
RP SEQUENCE FROM N.A.
RA DECKER K.F.;
RL SUBMITTED (OCT-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
RN [5]
RP SEQUENCE OF 1-231 FROM N.A.
RC TISSUE-TAIL;
RA KIRISITS M.J., VARDIMON D., KONZ H.W., GILL T.J. III;
RL SUBMITTED (JUN-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
CC - FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CONDITIONS.
CC - SUBUNIT: HOMOTRIMER.
CC - SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
EXTRACELLULAR SOLUBLE FORM.
CC - PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
PROTEOLYTIC PROCESSING.
CC - DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
AND MALNUTRITION.
CC - SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: X66539; G395370; -;
DR EMBL: L00981; G205254; -;
DR EMBL: D00475; G220921; -;
DR EMBL: AJ002278; E1154446; -;
DR EMBL: L19123; G310232; -;
DR PIR: J00029; J00029.
DR PIR: S21674; S21674.
DR PIR: J00868; J00868.
DR HSSP: P01375; 1TNE.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPEP 1 79
FT CHAIN 80 235 TUMOR NECROSIS FACTOR.

FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DISULFID 148 179 BY SIMILARITY.
 FT CARBOHYD 86 86 POTENTIAL.
 FT CONFLICT 39 39 I -> P (IN REF. 3 AND 4).
 FT CONFLICT 163 163 I -> T (IN REF. 3 AND 4).
 FT CONFLICT 202 202 F -> S (IN REF. 3 AND 4).
 SQ SEQUENCE 235 AA; 25806 MW; CIDFL87 CRC32;

Query Match 9.0%; Score 96; DB 1; Length 235;
 Best Local Similarity 27.5%; Pred. No. 0.044;
 Matches 58; Conservative 29; Mismatches 68; Indels 56; Gaps 15;

QY 22 GTGPGSNGEGYPMQSLPEQSSDLEAMENGRKRRAVLTKQKKQHS---VLHLVP 77
 DB 54 GVIGPKNK-EKFP-NGPLISSMA-----QTLTSSSGNSSDKRVAAHYVA 97
 QY 78 INATSKSDSYTEVMQPALRGGLAQGYGVR-----IDAGVYLXSVLPDVF 129
 DB 98 NHOAE-----OLEW--LSQRANALLANGMDLNQVLPADLYLYSVQLGOGGCP 149
 QY 129 --VFTMGQVVSREGQROE--TLFRCIRS--MPSHPDRA-----YNSQYAGVPHLHOG 177
 DB 150 DYVLLT--HYVSRAISYQEVSLLSAISKSPCKDTPEGALKEPTEPMYLGVFQLEKG 207
 QY 178 DIIS--VIIPRAKLNLSPHG-TFLGFVKL 205
 DB 208 DLISAENVLPK---YLDITSGQYVFVIAL 235

RESULT 2

TNFA_CAVPO STANDARD; PRT; 234 AA.

AC P51435;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS CAVIA PORCELLUS (GUINEA PIG).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; RODENTIA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-HARTLEY; TISSUE-LUNG;
 RA YUAN H.T., KELLY F.J., BINGLE C.D.;
 RL SUBMITTED (NOV-1995) TO EMBL/GENBANK/DBJ DATA BANKS.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-DUNKIN-HARTLEY;
 RA WHITE A.M., YOSHIMURA T., SMITH A.W., WESTWICK J., WATSON M.L.;
 RL SUBMITTED (NOV-1996) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY PROTEOLYTIC PROCESSING (BY SIMILARITY).
 CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH AND MALNUTRITION.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL: U39839; G1066112; -
 DR EMBL: U77036; G1679724; -
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.

FT PROPEP 1 79 BY SIMILARITY.
 FT CHAIN 80 234 TUMOR NECROSIS FACTOR.
 FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DISULFID 147 178 BY SIMILARITY.
 SQ SEQUENCE 234 AA; 25793 MW; 00160783 CRC32;

Query Match 9.0%; Score 95.5; DB 1; Length 234;
 Best Local Similarity 26.5%; Pred. No. 0.048;
 Matches 48; Conservative 22; Mismatches 62; Indels 49; Gaps 10;

QY 47 EAWENGRSRKRAVLTKQKKQHS---VLHVPINATSKSDSYTEVMQPALRGRC 102
 DB 62 EGFSSGPPFPPLAQTLTIRASQNDNDKPVAAHY-ANQAAE-----ELQW--LSKRANA 113
 QY 103 LQAGGYGVR-----IDAGVYLXSVLPDVFYTMQ-----VYREGQGRGT 147
 DB 114 LLANGKLSNQLVPSDGLYLYSVLFK-----GQCPSPYLLTHTVSRVAVSPEK 167
 QY 148 L-----FRCIRSPSPHPDRA--YNSQYAGVPHLHOGDILSVIIPRAKLNLSPHGT 198
 DB 168 VNLISATKSPCQKETPEGARKEPTEPMYLGVFQQLKGDRLS-----AEVNLPOYLD 220
 QY 199 F 199
 DB 221 F 221

RESULT 3

TNFA_PAPSP STANDARD; PRT; 233 AA.

AC P33620;
 DT 01-FEB-1994 (REL. 28, CREATED)
 DT 01-FEB-1994 (REL. 28, LAST SEQUENCE UPDATE)
 DT 01-FEB-1994 (REL. 28, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS PAPIO SP. (BABOON).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; PRIMATES.
 RN [1]
 RP SEQUENCE FROM N.A.
 RA SANJANMALA M., EDWARDS A.;
 RL SUBMITTED (SEP-1991) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN EXTRACELLULAR SOLUBLE FORM.
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY PROTEOLYTIC PROCESSING.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL: X62141; G38160; -
 DR PIR: S22052; S22052.
 DR HSSP: P01375; TNF.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR;
 KW MYRISTYLATION.
 FT PROPEP 1 76 BY SIMILARITY.
 FT CHAIN 77 233 TUMOR NECROSIS FACTOR.
 FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT LIPID 19 19 MYRISTATE (BY SIMILARITY).
 FT LIPID 20 20 MYRISTATE (BY SIMILARITY).
 FT DISULFID 145 177 BY SIMILARITY.
 SQ SEQUENCE 233 AA; 25557 MW; 7F049248 CRC32;

```

Query Match          8.7%; Score 92; DB 1; Length 233;
Best Local Similarity 27.2%; Pred. No. 0.1;
Matches 41; Conservative 23; Mismatches 51; Indels 36; Gaps 8;

QY      INATKDDSD--VFVMMOPAL-----RRGRLOAGGCVR-----IDAGVYLAYS 122
Db       77 VRSSRTFSDXRPVAVHVAVNPQAEGLQWLNFRANMLANGVELRNQLVWPSEGLIAYS 136
QY      123 QVLFPD-----VFETMGV-VSRREGOGROELTFR-CIRSNPSHDRA-YNISOYS 168
Db       137 QVLFKGQCSPSTHVLLTHTRISIANVSXTKVNLBLAISPCQRRETFEGEAAPWEPIYL 196
QY      169 AGVPHLHGODLITVIIPRAKLINSPGCTF 199
Db       197 GGVFOLERKGDRLS-----AEINLPDYLD F 220

RESULT   4
TNEA_HUMAN STANDARD; PRT; 233 AA.
AC        PO1375;
DT       21-JUL-1986 (REL. 01, CREATED)
DT       21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)
DT       15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE       TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN       TNFA.
OS       HOMO SAPIENS (HUMAN).
OC       EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
CC       EUTHERIA; PRIMATES.
[1]
RN       RP
RP       SEQUENCE FROM N.A.
RX       MEDLINE: 87217060.
RA       NEOSPASOV S.A., SHAKHOV A.N., TUTRESKAYA R.L., METT V.A.,
RA       AZIZOV M.M., GEORGIEV G.P., KOROBKO V.G., DOBRYNIN V.N.,
RA       FILIPOV S.A., BYSTROY N.S., BOLDYREVA E.F., CHUVPILO S.A.,
RA       CHIMAROV A.M., SHINGAROVA L.N., OVCHINNIKOV Y.A.;
RL       COLD SPRING HARB. SYMP. QUANT. BIOL. 51:611-624(1986).
[2]
RN       RN
RP       SEQUENCE FROM N.A.
RX       MEDLINE: 85086244.
RA       PENNICA D., NEDWIN J., HAYFLICK J.S., SEEBURG P.H., DENYXCK R.,
RA       PALADINO M.A., KOHR W.J., AGGARWAL B.B., GOEDEL D.V.;
RL       NATURE 312:724-729(1984).
[3]
RN       RN
RP       SEQUENCE FROM N.A.
RX       MEDLINE: 85137898.
RA       SHIRAI T., YAMAGUCHI H., ITO H., TODD C.W., WALLACE R.B.;
RL       NATURE 313:803-806(1985).
[4]
RN       RN
RP       SEQUENCE FROM N.A.
RX       MEDLINE: 86016093.
RA       NEMIN G.E., NAVLOR S.L., SAKAGUCHI A.Y., SMITH D.H.,
RA       JARETT-NEDWIN J., PENNICA D., GOEDEL D.V., GRAY P.W.;
RL       NUCLEIC ACIDS RES. 13:6361-6373(1985).
[5]
RN       RN
RP       SEQUENCE FROM N.A.
RX       MEDLINE: 85142190.
RA       WANG A.M., CEASESEY A.A., LADNER M.B., LIN L.S., STRICKLER J.,
RA       VAN ARSDELL J.N., YAMAMOTO R., MARK D.F.;
RL       SCIENCE 228:149-154(1985).
[6]
RN       RN
RP       SEQUENCE FROM N.A.
RX       MEDLINE: 86030296.
RA       MAARMENOT A., FRANSSEN L., TAVERNIER J., DER HEUDEN J., TIZARD R.,
RA       KANASHIMA E., SHAH V., JOHNSON M.J., SEMON D., MOELLER R.,
RA       RUSSCHAERT M.R., VAN VLIER A., FLERS W.;
RL       EUR. J. BIOCHEM. 152:515-522(1985).
[7]
RN       RN
RP       SEQUENCE FROM N.A.
RX       MEDLINE: 93272029.
RA       IRIS F.J.M., BOUGUELERET L., PRIEUR S., CATERRINA D., PRINAS G.,
RA       PERROT V., JURKA J., RODRIGUEZ-TOME P., CLAVERIE J.-M., DAUSSET J.,
RA       COHEN D.;

```

| | | | |
|----|---|---|------------------|
| RL | NAT. | GENET. | 3:137-145(1993). |
| RN | [8] | X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS). | |
| RP | MEDLINE; | 89159409. | |
| RA | JONES E.Y., STUART D.I., WALKER N.P.; | | |
| RL | NATURE | 338:225-228(1989). | |
| RN | [9] | X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS). | |
| RP | MEDLINE; | 91193276. | |
| RA | JONES E.Y., STUART D.I., WALKER N.P.; | | |
| RL | J. CELL SCI. SUPPL. | 13:11-18(1990). | |
| RN | [10] | X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS). | |
| RP | MEDLINE; | 90008932. | |
| RA | ECK M.J., SPRANG S.R.; | | |
| RL | J. BIOL. CHEM. | 264:17595-17605(1989). | |
| RN | [11] | X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF MUTANT ARG-107. | |
| RP | MEDLINE; | 98147459. | |
| RA | REDD C., FU Z.Q., WU J., XUE Y.N., HARRISON R.W., CHEN M.J., WEBER I.T.; | | |
| RL | PROTEIN ENG. | 10:1101-1107(1997). | |
| RN | [12] | MUTAGENESIS. | |
| RP | MEDLINE; | 91184128. | |
| RA | OSTADE X.V., TAVERNIER J., PRANGE T., FIEBS W.; | | |
| RL | EMBO J. | 10:827-836(1991). | |
| RN | [13] | MARIPOYLATION. | |
| RP | MEDLINE; | 93018820. | |
| RA | STEVENSON F.T., BURSTEN S.L., LOCKSLEY R.M., LOVETT D.H.; | | |
| RL | J. EXP. MED. | 176:1053-1062(1992). | |
| CC | -1 FUNCTION: | TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF CERTAIN TUMOR CELL LINES. IT IS IMPLICATED IN THE INDUCTION OF CACHEXIA, IT IS A POTENT PROGEN CAUSING FEVER BY DIRECT ACTION OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN CONDITIONS. | |
| CC | -1 SUBUNIT: | HOMOTRIMER. | |
| CC | -1 SUBCELLULAR LOCATION: | TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN EXTRACELLULAR SOLUBLE FORM. | |
| CC | -1 PTM: | THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY PROTEOLYTIC PROCESSING. | |
| CC | -1 DISEASE: | CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH AND MALNUTRITION | |
| CC | -1 SIMILARITY: | BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY. | |
| DR | EMBL; | X02910; G37210; -. | |
| DR | EMBL; | M16441; G339741; -. | |
| DR | EMBL; | X01394; G37220; -. | |
| DR | EMBL; | M10988; G339738; -. | |
| DR | EMBL; | M26331; G339764; -. | |
| DR | EMBL; | Z15026; G37212; -. | |
| DR | PIR; | B23784; QWHUN. | |
| DR | PIR; | A44189; A44189. | |
| DR | PIR; | S36153; S36153. | |
| DR | PDB; | 1TNF: 15-JAN-91. | |
| DR | PDB; | 2TUN: 31-JAN-94. | |
| DR | PDB; | 1A8H; 17-JUN-98. | |
| DR | MM; | 191160; -. | |
| DR | PROSITE; | PS00251; TNF_1; 1. | |
| DR | PROSITE; | PS50049; TNF_2; 1. | |
| KW | CYTOKIN; CYTOKIN-LIKE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR; MYOXYLATION; 3D-STRUCTURE. | | |
| FT | PROPEP | 1 | 76 |
| FT | CHAIN | 77 | 233 |
| FT | TRANSMEM | 36 | 56 |
| FT | LIPID | 19 | 19 |
| FT | LIPID | 20 | 20 |
| FT | DISULFID | 145 | 177 |
| FT | MOTAGEN | 105 | 105 |
| FT | MOTAGEN | 108 | 108 |
| FT | | | |

FT MTAGEN 112 112 L->F: BIOLOGICALLY INACTIVE.
 FT MTAGEN 160 160 A->V: BIOLOGICALLY INACTIVE.
 FT MTAGEN 162 162 S->F: BIOLOGICALLY INACTIVE.
 FT MTAGEN 167 167 V->A,D: BIOLOGICALLY INACTIVE.
 FT MTAGEN 222 222 E->K: BIOLOGICALLY INACTIVE.
 FT CONFLICT 63 63 F -> S (IN REF. 5).
 FT TURN 85 86
 FT STRAND 89 94
 FT TURN 99 100
 FT STRAND 104 105
 FT TURN 109 110
 FT STRAND 112 113
 FT TURN 115 116
 FT STRAND 118 119
 FT STRAND 124 125
 FT STRAND 132 138
 FT STRAND 140 142
 FT STRAND 152 159
 FT STRAND 166 174
 FT STRAND 190 191
 FT STRAND 194 200
 FT STRAND 207 212
 FT HELIX 215 217
 FT STRAND 218 218
 FT TURN 224 225
 FT STRAND 226 232
 SQ SEQUENCE 233 AA; 25644 MW; 666D7069 CRC32;

Query Match 8.6%; Score 91.5; DB 1; Length 233;
 Best Local Similarity 28.3%; Pred. No. 0.11;
 Matches 39; Conservative 20; Mismatches 50; Indels 29; Gaps 7;

QY 78 INAKSDSD--VEVMQPAL-----RRKGLAQGCVR-----IDAGYLLYS 122
 DB 77 VRSSRRPDPKPAHVANPAEQOLOMLNRANALLANGELRDNLVPESEGILYLYS 136
 QY 123 QVLFQD-----VFTMGQV-VSREGQROETLFR--CIRSMPSHPRDA--YNSQYS 168
 DB 137 QVLFKGGGCPSTHVLTLTTRISRIASVYTKVNLSAIKSPQOREPEGAEMKPTPEYTL 196
 QY 169 AGVPHLHQGDILSVIIPR 186
 DB 197 GGVFQLEKGRDLSSAELNR 214

RESULT 5
 ID TNFA_RABIT STANDARD; PRT; 235 AA.
 AC P04924;
 DT 13-AUG-1987 (REL. 05, CREATED)
 DT 13-AUG-1987 (REL. 05, LAST SEQUENCE UPDATE)
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS ORYCTOLAGUS CUNICULUS (RABBIT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 91065534.
 RA SHAKHOV A.N., KUPRASH D.V., AZIZOV M.M., JONGENSEL C.V.,
 RA NEDOSPASOV S.A.; 215-221(1990).
 RL GENE 95:215-221(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86219172.
 RA ITO H., SHIRAI T., YAMAMOTO S., AKIRA M., KAWAHARA S., TODD C.W.,
 RA WALLACE R.B.; 5:157-165(1986).
 RL DNA 5:157-165(1986).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE; 86219711.

RA ITO H., YAMAMOTO S., KURODA S., SAKAMOTO H., KAJIHARA J., KIYOTA T.,
 RA HAYASHI H., KATO M., SEKO M.;
 RL DNA 5:149-156(1986).
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
 CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
 CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CC CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 CC EXTRACELLULAR SOLUBLE FORM.
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 CC PROTEOLYTIC PROCESSING.
 CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
 CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
 CC AND MALNUTRITION.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL; M12845; G165760; -;
 DR EMBL; M12846; G165753; -;
 DR EMBL; M60340; G165756; -;
 DR PIR; A25451; A25451.
 DR PIR; A25454; A25454.
 DR PIR; J50727; J50727.
 DR HSSP; P01375; 1TNF.
 DR PROSITE; PS00251; TNF_1; 1.
 DR PROSITE; PS50049; TNF_2; 1.
 KW CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT PROPEP 1 79
 FT CHAIN 80 235 TUMOR NECROSIS FACTOR.
 FT TRANSSEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 FT DISULFID 148 179 BY SIMILARITY.
 FT CONFLICT 63 63 MISSING (IN REF. 3).
 SQ SEQUENCE 235 AA; 25816 MW; 1BC5C08 CRC32;

Query Match 8.5%; Score 90.5; DB 1; Length 235;
 Best Local Similarity 24.1%; Pred. No. 0.14;
 Matches 48; Conservative 28; Mismatches 54; Indels 69; Gaps 12;

QY 59 RAVLTOKOKROHSLVHL-----VPINATSK--DSDVTEVMQPAL-----RRGR 101
 DB 54 RVIGPQEEGSPNNLHLYNPVQGMVTLRSASRALSDKPLAHVYANPVEGQLQWLSGRAN 113
 QY 102 GLQAGTGVRIQD-----AGVLLYSQVLFQDVFTMGQVVSREGQ----- 144
 DB 114 ALLAN--GMKLTDLQVLVPADGLYLYSQVLF-----SQGCRSYVLLTHTV 158
 QY 144 -----ROETLFRCTIRSMPSH--PDRA-----YNSQYSAGVPHLHQGDILSVIIPRA 187
 DB 159 SRFAVSTPNKYNLSSAKS--PCHRETPDEAPMAMVPEYTLGGVFQLEKGRDLSTEVENP 217
 QY 188 RAKLNSPHG-TFLGFYKL 205
 DB 218 E-YIDLAESGQVYRGITALL 235

RESULT 6
 ID TNFA_CEREL STANDARD; PRT; 229 AA.
 AC P51743;
 DT 01-OCT-1996 (REL. 34, CREATED)
 DT 01-OCT-1996 (REL. 34, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN) (FRAGMENT).
 GN TNFA.
 OS CERVUS ELAPHUS (RED DEER).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA LOCKHART E.A.;

```

RL SUBMITTED (SEP-1994) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM (BY SIMILARITY).
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING (BY SIMILARITY).
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: U14683; G540279; -.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KM CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT NON_TER 1
FT PROPEP 1
FT CHAIN 73 BY SIMILARITY.
FT TRANSEM 74 229 TUMOR NECROSIS FACTOR.
FT DISULFID 31 51 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DISULFID 141 173 BY SIMILARITY.
SQ SEQUENCE 229 AA; 24987 MW; 4FD25D97 CRC32;

Query Match 8.3%; Score 88; DB 1; Length 229;
Best Local Similarity 28.8%; Pred. No. 0.24;
Matches 47; Conservative 24; Mismatches 40; Indels 52; Gaps 13;

QY 61 VLTKQKKQKHS---VLHLP-INAISKDDSDVTEVMQPALRGRGLQAQGYRIOD- 115
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DQ 70 VQTLRSSQASINKPVAVHVNINNOG-----QLWLMDSC--ANALMAN--GVKLEDN 118
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 115 -----AGVYLISQVLYLQD-----VFTMGQV-VSRGGQROETLFRCLSPMSH 158
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DQ 119 QLVPTDLYLISQVLYLQDSCPSPTPLPLHTHTISRIASVYQ--TKVNIISAIKS-PCH 174
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

QY 159 ---PDRA-----YNSQYSGVPHLHQGDLISVILIRAKAKML 193
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
DQ 175 RHPMAEAKPWEPIYQGVGFLEKGRRLS-----AEIWL 210
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

RESULT 7
ID TNFA_MOUSE STANDARD: PRT: 235 AA.
AC P06804;
DT 01-JAN-1988 (REL. 06, CREATED)
DT 01-MAR-1989 (REL. 10, LAST SEQUENCE UPDATE)
DT 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
DN TNFA OR TNF.
OS MUS MUSCULUS (MOUSE).
OC EUAROTIA; METAEOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; RODENTIA.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88224564.
RA SHIRAI T., SHIMIZU N., SHIOJIRI S., HORIGUCHI S., ITO H.;
RL DNA 7:193-201(1988).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE: 85298296.
RA PENNICA D., HAYFLICK J.S., BRINGMAN T.S., PALLADIO M.A.,
RA GOSDEDEL D.V.;
RL PROC. NATL. ACAD. SCI. U.S.A. 82:6060-6064(1985).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE: 86149365.
RA CARUT D., BEUTLER B., HARTOG K., THAYER R., BROWN-SHIMER S.,

```

```

RA CERAMI A.;
RL PROC. NATL. ACAD. SCI. U.S.A. 83:1670-1674(1986).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE: 85242112.
RA FRANSSEN L., MULLER R., MARMENOUT A., TAVERNIER J., VAN DER HEYDEN J.,
RA KAWASHIMA E., COLLETT A., TIZARD R., VAN HEUVERSWIN H., VAN VLIET A.,
RA RUTSCHAERT M.-R., FIEBS W.;
RL NUCLEIC ACIDS RES. 13:4417-4429(1985).
RN [5]
RP SEQUENCE FROM N.A.
RX MEDLINE: 87298639.
RA SHAKHOV A.N., NEDOSPASOV S.A.;
RL BIOORG. KHIM. 13:701-705(1987).
RN [6]
RP SEQUENCE FROM N.A.
RX MEDLINE: 88067722.
RA SEMON D., KAWASHIMA E., JONGENEEL C.V., SHAKHOV A.N., NEDOSPASOV S.A.;
RL NUCLEIC ACIDS RES. 15:9083-9084(1987).
RN [7]
RP SEQUENCE FROM N.A.
RX STRAIN=CTS, AND NOD;
RX MEDLINE: 96013654.
RA IREGAMI H., MAKINO S., YAMATO E., KAWAGUCHI Y., UEDA H., SAKAMOTO T.,
RA TAKEKAWA K., OGIHARA T.;
RL J. CLIN. INVEST. 96:1936-1942(1995).
RN [8]
RP SEQUENCE OF 80-99.
RX MEDLINE: 91097531.
RA SHERRY B., JUC D.-M., ZENTELLA A., CERAMI A.;
RL BIOCHEM. BIOPHYS. RES. COMMUN. 173:1072-1078(1990).
RN [9]
RP SEQUENCE OF 70-87.
RX MEDLINE: 89380231.
RA CSEH K., BEUTLER B.;
RL J. BIOL. CHEM. 264:16256-16260(1989).
RN [10]
RP IDENTIFICATION OF MEMBRANE-BOUND FORM.
RX MEDLINE: 88165056.
RA KRIEGLER M., PEREZ X., DEFRAY K., ALBERT I., LU S.D.;
RL CELL 53:45-53(1988).
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES, IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: U06950; G495468; -.
DR EMBL: M13049; G302083; -.
DR EMBL: M11731; G302085; -.
DR EMBL: Y00467; G54832; -.
DR EMBL: X02611; G54845; -.
DR EMBL: M20155; G202093; ALT_SEQ.
DR EMBL: M38296; G202087; -.
DR EMBL: D84196; D1020271; -.
DR EMBL: D84194; D1020271; JOINED.
DR EMBL: D84195; D1020271; JOINED.
DR EMBL: D84197; D1020272; -.
DR EMBL: D84198; D1020272; JOINED.
DR PIR: A23127; OWMNS.
DR PIR: A22908; A22908.
DR PIR: A23164; A23164.

```

DR PIR: A27303; A27303.
 DR PIR: A34251; A34251.
 DR PIR: S03791; S03791.
 DR HSSP: P01375; 1TNF.
 DR MGI: MGI:104798; 1TNF.
 DR PROSITE: PS00251; 1TNF_1; 1.
 DR PROSITE: PS0049; 1TNF_2; 1.
 DR CYTOKINE; CYTOKIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT CHAIN 1 79
 FT PROPEP 1 79
 FT TRANSMEM 36 235
 FT DISULFID 148 179
 FT CARBOHYD 86 86
 FT CONFLICT 231 231
 FT SEQUENCE 235 AA; 25895 MW; 7FDE3C71 CRC32; G -> R (IN REF. 3 AND 4).

Query Match 8.2%; Score 87.5; DB 1; Length 235;
 Best Local Similarity 27.7%; Pred. No. 0.28;
 Matches 56; Conservative 22; Mismatches 67; Indels 57; Gaps 13;

QY 22 GTCGSPQNEGYPMWQSLPEQSSDALEAMENGERSKRRRAVLTKOKKSHS---VLAHVP 77
 Db 54 GVIQV-QRDEKEP-NGPLPISMA-----QTLTLRSSSSONSQKPAHAYVA 97
 QY 78 INATSKDSDVTEVMQPALRRGRLQAOGYVR-----IQDAGVYLKSOVLFQD---- 129
 Db 98 NHOVE-----QLEM--LSQRANALLANGMDLKNQVVPADGLVLYSOVLFKQGC 149
 QY 129 --VFTMGVQVNSREGGRDE--TLFRGIRS--MPSHPDRA---YNSQYSAQVPHHOG 177
 Db 150 DYVLTL--HTVSFAISYQEKVILSAVSPCKDPPEGAELKPMYEPILYGVFQLENG 207
 QY 178 DILSVIIPRARKLNLSPHGT 199
 Db 208 DQLS-----AEVNLPKYLD 222

RESULT 8

TNFA_BOVIN ID TNFA_BOVIN STANDARD; PRT; 233 AA.
 AC 006599;
 DT 01-JUN-1994 (REL. 29, CREATED)
 DT 01-JUN-1994 (REL. 29, LAST SEQUENCE UPDATE)
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA OR TNF.
 OS BOS TAURUS (BOVINE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; ARTIODACTYLA.
 RN 11
 RN RP SEQUENCE FROM N.A.
 RX MEDLINE; 94083525.
 RA CLUDIS I., CLEUTER Y., KETTMANN R., BURNY A., DROOGMANS L.;
 RL CYTOKINE 5:336-341(1993).
 RN 12
 RP SEQUENCE OF 50-233 FROM N.A.
 RC TISSUE-BLOOD;
 RX MEDLINE; 96006582.
 RA MERTENS B.E.L.C., MORIURI M., GAUDULIS L.;
 RL IMMUNOGENETICS 42:430-431(1995).
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
 WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
 CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 CC PROTEOLYTIC PROCESSING.

CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
 CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
 CC AND MALNUTRITION.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL: Z14137; G798; -
 DR EMBL: Z48808; G755702; -
 DR PIR: S24642; S24642.
 DR HSSP: P01375; 1TNF.
 DR PROSITE: PS00251; 1TNF_1; 1.
 DR PROSITE: PS0049; 1TNF_2; 1.
 DR CYTOKINE; CYTOKIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT CHAIN 1 77
 FT PROPEP 1 77
 FT TRANSMEM 36 233
 FT DISULFID 145 177
 FT CONFLICT 62 62
 FT SEQUENCE 233 AA; 25439 MW; 207AFB2E CRC32; E -> EQ (IN REF. 2).

Query Match 8.2%; Score 87; DB 1; Length 233;
 Best Local Similarity 25.5%; Pred. No. 0.3;
 Matches 51; Conservative 24; Mismatches 55; Indels 70; Gaps 13;

QY 19 RLQGTGSPQNEGYPMWQSLPEQSSDALEAMENGERSKRRRAVLTKOKKSHVLAHVP 78
 Db 60 REESGPGPSINS---PLVQTLRSSQA-----SNKPAHAYV-- 94
 QY 79 NATSKDSDVTEVMQPALRRGRLQAOGYVRID-----AGVYLKSOVLFQD---- 129
 Db 94 ---ADINSPGOLRWMDSY--ANALMAN--GVKLEDNQVVPADGLVLYSOVLFKQGC 145
 QY 129 ---VFETMGVQVNSREGGRDE--TLFRGIRS--MPSHPDRA---YNSQYSAQVPH 173
 Db 146 PSTPLFLHTTISRIVSTQ--TKVNIILSAIKS-PCHRETPEMAEKPMYEPITQGVFQ 201
 QY 174 LHQGDILSVIIPRARKLN 193
 Db 202 LEKGDRLS-----AEINL 214

RESULT 9

TNFA_FELCA ID TNFA_FELCA STANDARD; PRT; 233 AA.
 AC P19101;
 DT 01-NOV-1990 (REL. 16, CREATED)
 DT 01-NOV-1990 (REL. 16, LAST SEQUENCE UPDATE)
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS FELIS SILVESTRIS CATUS (CAT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 OC EUTHERIA; CARNIVORA.
 RN 11
 RN RP SEQUENCE FROM N.A.
 RX TISSUE-BLOOD;
 RX MEDLINE; 91016860.
 RA MCGRAW R.A., COFFEE B.W., OTTO C.M., DREWS R.T., RAWLINGS C.A.;
 RL NUCLEIC ACIDS RES. 18:5563-5563(1990).
 RN 12
 RP SEQUENCE FROM N.A.
 RC TISSUE-BONE MARROW;
 RA DANIEL S.L., BRENNER C.A., LEGENDRE A.M., SOLOMAN A., ROUSE B.T.;
 RL SUBMITTED (XX-1993) TO EMBL/GENBANK/DBJ DATA BANKS.
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
 WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CERTAIN TUMOR CELL LINES, IT IS IMPLICATED IN THE INDUCTION OF
 CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 OR BY STIMULATION OF INTERLEUKIN 1 SECRETION, IT CAN STIMULATE
 CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 CC EXTRACELLULAR SOLUBLE FORM.


```
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION. AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC EMBL: X54000; G295777; -.
CC DR EMBL: M92061; G403366; -.
CC DR PIR: S11688; S11688.
CC DR HSSP: P01375; TNF.
CC DR PROSITE: PS00251; TNF_1; 1.
CC DR PROSITE: PS50049; TNF_2; 1.
CC KM CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
CC FT PROPEP 1 76
CC FT CHAIN 77 233 TUMOR NECROSIS FACTOR.
CC FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
CC FT DISULFID 145 177 BY SIMILARITY.
CC FT CONFLICT 28 28 G -> R (IN REF. 2).
CC FT CONFLICT 104 104 R -> W (IN REF. 2).
CC FT CONFLICT 151 151 L -> H (IN REF. 2).
CC FT CONFLICT 155 155 A -> T (IN REF. 2).
CC FT CONFLICT 210 210 T -> A (IN REF. 2).
CC SQ SEQUENCE 233 AA; 25322 MW; 1A5651A6 CRC32;

Query Match 8.2%; Score 86.5; DB 1; Length 233;
Best Local Similarity 24.4%; Pred. No. 0.34;
Matches 43; Conservative 22; Mismatches 76; Indels 35; Gaps 8;

QY 64 QKQKQSHVLYLPINAT-----SKDDSDVTEVMQPAL-----RRRGIGQAGY- 109
DB 59 QREELRHGDLINLPOTLSSSRTPSDKPYAHVYANPEEGOLQRLSRANALANGVE 118
QY 109 ----GVRIQDAGVLYLYSYQLF-----QDVTFTMGQVYSEGGGROETL-----FRCI 152
DB 119 LFDNQKVPDGLYLYTSQVLFEGGCGPSTHVLTHAISFVSYQTKVNLSSKSPQ 178
QY 153 RSMPSHPDRA--YNSOYSAQVPHHOGDILSVIIPARAKLNSPHG-ITLGFVKL 205
DB 179 RETPEGAENAPWYEPITLIGVFOLEKGRDLSTEL-NLPAYLDFAESGOVYFGIATL 233

RESULT 10
TNFA_MACFA STANDARD; PRT; 233 AA.
AC P79337;
DT 15-JUL-1998 (REL. 36, CREATED)
DT 15-JUL-1998 (REL. 36, LAST SEQUENCE UPDATE)
DE 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS MACACA FASCICULARIS (CRAB EATING MACAQUE) (CYNOMOLGUS MONKEY).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LYMPHOCYTES.
RA TATSUMI M.;
RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES. IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION. IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
```

```
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC DR EMBL: AB000513; G1794149; -.
CC DR PROSITE: PS00251; TNF_1; 1.
CC DR PROSITE: PS50049; TNF_2; 1.
CC KM CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
CC FT PROPEP 1 76
CC FT CHAIN 77 233 TUMOR NECROSIS FACTOR.
CC FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
CC FT DISULFID 145 177 BY SIMILARITY.
CC SQ SEQUENCE 233 AA; 25558 MW; 856FC354 CRC32;

Query Match 8.1%; Score 86; DB 1; Length 233;
Best Local Similarity 26.5%; Pred. No. 0.38;
Matches 40; Conservative 23; Mismatches 52; Indels 36; Gaps 8;

QY 78 INATSDSD--VTEVMQPAL-----RRRGIGQAGYVR-----IQDAGVLYLS 122
DB 77 VRSSSRTPSDKPYAHVYANPEEGOLQWLNRANALVANGVELTDNQLVPSGSLYLS 136
QY 123 QVLFQD-----VFTMGQV-VEREGGROETLFR--CIRSMPSHPDRA--YNSOYS 168
DB 137 QVLFKQGGCPSNHVLTHTSRITAVSQTKVNLSSKSPQNETPEGAARKWYEPITL 196
QY 169 AGVPHLHOGDILSVIIPARAKLNSPHGTF 199
DB 197 GGVFOLEKGRDLS-----AEINLPDYLDLF 220

RESULT 11
TNFA_MACMU STANDARD; PRT; 233 AA.
AC P48094;
DT 01-FEB-1996 (REL. 33, CREATED)
DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)
DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
GN TNFA.
OS MACACA MULATTA (RHESUS MACAQUE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
OC EUTHERIA; PRIMATES.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 96003435.
RA VILINGER F.J.; BRAR S.S.; MAYNE A.E.; CHIKKALA N.; ANSARI A.A.;
RL J. IMMUNOL. 155:3946-3954(1995).
CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
CC WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
CC CERTAIN TUMOR CELL LINES. IT IS IMPLICATED IN THE INDUCTION OF
CC CACHEXIA, IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
CC OR BY STIMULATION OF INTERLEUKIN 1 SECRETION. IT CAN STIMULATE
CC CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
CC CONDITIONS.
CC -1- SUBUNIT: HOMOTRIMER.
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
CC EXTRACELLULAR SOLUBLE FORM.
CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
CC PROTEOLYTIC PROCESSING.
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC DR EMBL: U19850; G644818; -.
CC DR PROSITE: PS00251; TNF_1; 1.
CC DR PROSITE: PS50049; TNF_2; 1.
CC KM CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
CC FT PROPEP 1 76
CC FT CHAIN 77 233 TUMOR NECROSIS FACTOR.
CC FT TRANSMEM 36 56 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
CC FT DISULFID 145 177 BY SIMILARITY.
CC SQ SEQUENCE 233 AA; 25630 MW; 146A1B48 CRC32;
```

Query Match 8.1%; Score 86; DB 1; Length 233;
 Best Local Similarity 26.5%; Pred. No. 0.38;
 Matches 40; Conservative 23; Mismatches 52; Indels 36; Gaps 8;

QY 78 INATSKDSD--VTEVMQPAL-----RGRGLQAGYGVV----TODAGVYLLYS 122
 DB 77 VRSSRRPSPDKPAHVANPQAEQOLWLNRRANALLANGVELDNLQVPSGGLYIYS 136
 QY 123 QVLEFD-----VFTMQGV-VSREGGQRETLR--CIRSMPSHPRA--YNSQYS 168
 DB 137 QVLEFGGCGPSNHVLLHTHTISRLAVSYQTKVNULLSAIKSPCQRETPGAEAKPWEPYTL 196
 QY 169 AGVPHLHOGDILSVITPRRAKLNLSPHGTF 199
 DB 197 GGVQLEKGRDLS-----AEINLPDYLD 220

RESULT 12
 TNFB_RABIT
 ID TNFB_RABIT STANDARD; PRT; 197 AA.
 AC P10154;
 DT 01-MAR-1989 (REL. 10, CREATED)
 DT 01-AUG-1991 (REL. 19, LAST SEQUENCE UPDATE)
 DE LYMPHOTOXIN-ALPHA PRECURSOR (LT-ALPHA) (TNF-BETA).
 GN LTA OR TNFB.
 OS ORYCTOLAGUS CUNICULUS (RABBIT).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; LAGOMORPHA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 91065534.
 RA SHAKHOV A.N., KUPRASH D.V., AZIZOV M.M., JONGENEEL C.V.,
 RA NEDOSPASOV S.A.;
 RL GENE 95:215-221(1990).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 90220566.
 RA SHAKHOV A.N., KUPRASH D.V., TURETSKAYA R.L., AZIZOV M.M.,
 RA ANDREYEVA A.V., NEDOSPASOV S.A.;
 RL MOL. BIOL. (MOSC) 23:1743-1750(1989).
 CC -1- FUNCTION: LYMPHOTOXIN IS A CYTOKINE PRODUCED BY LYMPHOCYTES WHICH
 IS CYTOTOXIC FOR A WIDE RANGE OF TUMOR CELLS IN VITRO AND IN VIVO.
 CC -1- SUBUNIT: HETEROTRIMER OF EITHER TWO BETA AND ONE ALPHA SUBUNITS
 OR (LESS PREVALENT) TWO ALPHA AND ONE BETA SUBUNITS.
 CC -1- SUBCELLULAR LOCATION: SECRETED.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL: X55745; G297168; ALT_SEQ.
 DR EMBL: M60340; G165755; -.
 DR EMBL: M60341; G165758; -.
 DR PIR: JH0309; JH0309.
 DR PIR: PNO098; PNO098.
 DR HSSP: P01375; 1TNF.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 KM CYTOKINE; GLYCOPROTEIN; CYTOTOXIN; SIGNAL.
 FT SIGNAL 1 26
 FT CHAIN 27 197 LYMPHOTOXIN-ALPHA.
 FT CARBOHYD 88 88 POTENTIAL.
 SO SEQUENCE 197 AA; 21126 MW; 38A27DB CRC32;

Query Match 8.1%; Score 86; DB 1; Length 197;
 Best Local Similarity 22.8%; Pred. No. 0.3;
 Matches 49; Conservative 19; Mismatches 67; Indels 80; Gaps 10;

QY 26 PSQNGEGYPMOSLPEQSSDALEAWENGSRKRAVLTQKKQHSYV---HLVPINAT 81
 DB 20 PPRAGQLPFAEFP-----SAARNAQORLQAHFHSHTLKPAHLY----- 61
 QY 82 SKDDSDVTEVMQPALRG---RGLQAGYGVRIODAGVYLLYSQVLFQDVTFTMGQVVS 138

DB 61 -GDPSQDSLWRANDRAFLRHGFLSNNSLLVSPSSGLYFYVSQVVE----- 108
 QY 139 REGGROETLFRCL-RKMP-----SHPDRAVNSQYSAGVPHL----- 176
 DB 108 -SGEG-----CSKAVPPLYLAHEVQLFSSQYSFHVBLLSNOKSVCPDGPQVWSV 159
 QY 176 -----OGDILSVITPRRAKLNLSPHGTF 201
 DB 160 YQAGVFLITQGDQLSTHTD-GIAHLILSPSSVFFG 193

RESULT 13
 TNFA_HORSE
 ID TNFA_HORSE STANDARD; PRT; 234 AA.
 AC P29553;
 DT 01-APR-1993 (REL. 25, CREATED)
 DT 01-APR-1993 (REL. 25, LAST SEQUENCE UPDATE)
 DE TUMOR NECROSIS FACTOR PRECURSOR (TNF-ALPHA) (CACHECTIN).
 GN TNFA.
 OS EQUUS CABALLUS (HORSE).
 OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; TETRAPODA; MAMMALIA;
 CC EUTHERIA; PERISSODACTYLA.
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE: 92084125.
 RA ST X., MORRIS D.D., MCGRAW R.A.;
 RL GENE 107:319-321(1991).
 CC -1- FUNCTION: TNF IS MAINLY SECRETED BY MACROPHAGES. IT IS A CYTOKINE
 WITH A WIDE VARIETY OF FUNCTIONS: IT CAN CAUSE CYTOLYSIS OF
 CERTAIN TUMOR CELL LINES. IT IS IMPLICATED IN THE INDUCTION OF
 CACHECTIN. IT IS A POTENT PYROGEN CAUSING FEVER BY DIRECT ACTION
 OR BY STIMULATION OF INTERLEUKIN 1 SECRETION. IT CAN STIMULATE
 CELL PROLIFERATION AND INDUCE CELL DIFFERENTIATION UNDER CERTAIN
 CONDITIONS.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 CC EXTRACELLULAR SOLUBLE FORM.
 CC -1- PTM: THE SOLUBLE FORM DERIVES FROM THE MEMBRANE FORM BY
 CC PROTEOLYTIC PROCESSING.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 DR EMBL: M64087; G164245; -.
 DR PIR: J01344; J01344.
 DR HSSP: P01375; 1TNF.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 KM CYTOKINE; CYTOTOXIN; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
 FT PROPEP 1 77
 FT CHAIN 78 234
 FT TRANSMEM 36 56 TUMOR NECROSIS FACTOR.
 FT DISULFID 146 178 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
 SO SEQUENCE 234 AA; 25469 MW; CD20BEE CRC32;

Query Match 8.1%; Score 85.5; DB 1; Length 234;
 Best Local Similarity 27.3%; Pred. No. 0.42; DB 1; Length 234;
 Matches 41; Conservative 14; Mismatches 38; Indels 57; Gaps 8;

QY 78 INATSKDSD--VTEVMQPALR-----RGRGLQAGYGVRIOD-----AGVYLLYS 122
 DB 78 LRSSRRPSPDKPAHVANPQAEQOLWLSGRANALLANGVELDNLQVPLDGLYIYS 137
 QY 123 QVLEFD-----VFTMQGV-VSREGGQRETLR--CIRSMPSHPRA--YNSQYS 168
 DB 138 QVLEFD-----VFTMQGV-VSREGGQRETLR--CIRSMPSHPRA--YNSQYS 168
 QY 160 DRA-----YNSQYSAGVPHLHOGDILSVII 184
 DB 184 EQAEAKPWEPYTLIGGVQLEKGDQLSAET 213

RESULT 14
 TNFB_HUMAN

CC PROTEOLYTIC PROCESSING (BY SIMILARITY).
CC -1- DISEASE: CACHEXIA ACCOMPANIES A VARIETY OF DISEASES, INCLUDING
CC CANCER AND INFECTION, AND IS CHARACTERIZED BY GENERAL ILL HEALTH
CC AND MALNUTRITION.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
DR EMBL: X94932; E222187; -.
DR EMBL: S74068; G802045; -.
DR EMBL: Z70046; E228213; -.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KM CYTOKINE: CYTOKINE; TRANSMEMBRANE; GLYCOPROTEIN; SIGNAL-ANCHOR.
FT PROPER 1 76
FT CHAIN 77 233
FT TRANSMEM 36 56
FT DISULFID 145 177
FT CONFLICT 59 60
FT CONFLICT 66 66
FT CONFLICT 74 74
FT CONFLICT 111 111
FT CONFLICT 116 116
FT CONFLICT 134 135
SQ SEQUENCE 233 AA; 25447 MW; BCA64FA0 CRC32;
POTENTIAL.
TUMOR NECROSIS FACTOR.
SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
BY SIMILARITY.
OR -> PE (IN REF. 2).
G -> C (IN REF. 2).
A -> V (IN REF. 3).
A -> D (IN REF. 2).
G -> D (IN REF. 2).
IY -> DS (IN REF. 2).

Query Match 7.9%; Score 83.5; DB 1; Length 233;
Best Local Similarity 24.7%; Pred. No. 0.65;
Matches 43; Conservative 26; Mismatches 62; Indels 43; Gaps 9;
QY 64 QOKOKSHVHLIVINATSK-----DDSDYTEVMQPAL-----RRGRGLQAQGYG 109
Db 59 QRELPNGLOLISPLAQTIVKSSRTPSDKPYAHVYANPEAGQLQWLRRANALLANGVE 118
QY 110 VR-----IODAGVYLXSYQLFOD-----VFTMGQ-VVREGOGROETLFR--CI 152
Db 119 LITDNLIVPSDGLILYSQVLFKRGCGCPSTHYVLTHITISRAVSYQTKVNLISAIKSPCQ 178
QY 153 RSMSPHPDRA--YNSQYSAGVPHLHQGDILSVIIPRAKILNLSPHGTFLEGVFR 204
Db 179 RETPEGTETAKPWYEPYILGVFQLEKGBRLS-----AEINLP--NYLDFAE 222

Search completed: May 19, 1999, 14:43:40
Job time: 47 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 14, 1999, 21:39:58 ; Search time 48.52 Seconds

(without alignments)
233.093 Million cell updates/sec

Title: US-09-212-270-5

Perfect score: 1061

Sequence: 1 MVLQQTQLQSLRREVSR.....RAKLNLSPHGFLGFVKL 205

Scoring table: BLOSUM62

Searched: 180763 seqs, 55169189 residues

Database :

SPTREMBL_8:*
1: sp_fungi:*
2: sp_human:*
3: sp_invertebrate:*
4: sp_mammal:*
5: sp_mhc:*
6: sp_organelle:*
7: sp_phage:*
8: sp_plant:*
9: sp_bacteria:*
10: sp_rodent:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_archaea:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|--------|--------------------|
| 1 | 1031 | 97.2 | 250 | 2 | 075888 | 075888 homo sapien |
| 2 | 107.5 | 10.1 | 391 | 2 | 075910 | 075910 homo sapien |
| 3 | 90.5 | 8.5 | 157 | 2 | 043647 | 043647 homo sapien |
| 4 | 89.5 | 8.4 | 377 | 10 | 054693 | 054693 mus musculi |
| 5 | 87.5 | 8.2 | 583 | 8 | 043143 | 043143 lycopersico |
| 6 | 87 | 8.2 | 233 | 4 | 018779 | 018779 bos taurus |
| 7 | 87 | 8.2 | 1726 | 10 | 062383 | 062383 mus musculi |
| 8 | 86 | 8.1 | 233 | 4 | 07510 | 07510 papio hamad |
| 9 | 85.5 | 8.1 | 1375 | 3 | 094537 | 094537 dirosophila |
| 10 | 85.5 | 8.1 | 1526 | 3 | 094538 | 094538 dirosophila |
| 11 | 85.5 | 8.1 | 364 | 10 | 054910 | 054910 mus musculi |
| 12 | 85.5 | 8.1 | 319 | 12 | 093529 | 093529 xenopus lae |
| 13 | 83 | 7.8 | 754 | 11 | 089250 | 089250 soil-borne |
| 14 | 82 | 7.7 | 232 | 10 | 035853 | 035853 mus musculi |
| 15 | 81 | 7.6 | 392 | 5 | 031277 | 031277 rattus norv |
| 16 | 81 | 7.6 | 233 | 10 | 035734 | 035734 marmota mon |
| 17 | 80.5 | 7.6 | 377 | 5 | 031279 | 031279 rattus norv |
| 18 | 80.5 | 7.6 | 357 | 9 | 083906 | 083906 treponema p |
| 19 | 80 | 7.5 | 1603 | 2 | 015737 | 015737 homo sapien |
| 20 | 80 | 7.5 | 2957 | 3 | 061845 | 061845 caenorhabd1 |
| 21 | 80 | 7.5 | 513 | 9 | 054427 | 054427 salmonella |
| 22 | 80 | 7.5 | 707 | 9 | 005415 | 005415 clostridium |
| 23 | 79.5 | 7.5 | 234 | 4 | 028320 | 028320 capra hircu |
| 24 | 79 | 7.4 | 319 | 9 | 048367 | 048367 enterococcu |
| 25 | 79 | 7.4 | 307 | 9 | 087203 | 087203 lactococcu |
| 26 | 79 | 7.4 | 2629 | 10 | P97499 | P97499 mus musculi |
| 27 | 79 | 7.4 | 1888 | 10 | 088466 | 088466 mus musculi |
| 28 | 78.5 | 7.4 | 216 | 10 | 070332 | 070332 mesocricetu |
| 29 | 78.5 | 7.4 | 842 | 10 | 008874 | 008874 rattus norv |

ALIGNMENTS

| RESULT | 1 | PRELIMINARY; | PRT; | 250 AA. |
|---------------------------|--|--------------|------|---------|
| ID | 075888 | | | |
| AC | 075888 | | | |
| DT | 01-NOV-1998 (TREMBLREL. 08, CREATED) | | | |
| DT | 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE) | | | |
| DE | PROLIFERATION INDUCING LIGAND APRIL. | | | |
| OS | HOMO SAPIENS (HUMAN). | | | |
| OC | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; | | | |
| OC | PRIMATES; CATARRHINI; HOMINIDAE; HOMO. | | | |
| RN | [1] | | | |
| RP | SEQUENCE FROM N.A. | | | |
| RC | TISSUE-UTERUS. | | | |
| RX | MEDLINE; 98416181. | | | |
| RA | HAHNE M., KATZACKA T., SCHROETER M., HOFMANN K., IRMLER M., | | | |
| RA | BODMER J.-L., SCHNEIDER P., BORNAND T., HOLLER N., FRENCH L.E., | | | |
| RA | SORDAT B., RIMOLDI D., TSCHOOP J. | | | |
| RT | "APRIL, a New Ligand of the Tumor Necrosis Factor Family, Stimulates | | | |
| RT | Tumor Cell Growth." | | | |
| RL | J. EXP. MED. 188:1185-1190(1998). | | | |
| DR | EMBL; AF046888; G3650492; - | | | |
| SQ | SEQUENCE 250 AA; 27433 MW; 48C59C35 CRC32; | | | |
| Query Match | 97.28; Score 1031; DB 2; Length 250; | | | |
| Best Local Similarity | 98.08; Pred. No. 3.7e-85; | | | |
| Matches 201; Conservative | 0; Mismatches 4; Indels 0; Gaps 0; | | | |
| QY | 1 | | | |
| DB | 46 | | | |
| QY | 61 | | | |
| DB | 106 | | | |
| QY | 121 | | | |
| DB | 166 | | | |
| QY | 181 | | | |
| DB | 226 | | | |
| RESULT | 2 | | | |
| ID | 075910 | | | |
| AC | 075910 | | | |
| DT | 01-NOV-1998 (TREMBLREL. 08, CREATED) | | | |
| DT | 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE) | | | |

DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE ECTODYSPLASIN-A EDA ISOFORM II.
GN ED1.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA.
OC PRIMATES; CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE: 98349961.
RA MONREAL A.W., ZONANA J., FERGUSON B.;
RT "Identification of a new splice form of the ED1 gene permits
RT detection of nearly all x-linked hypohidrotic ectodermal dysplasia
RT mutations.";
RL AM. J. HUM. GENET. 63:380-389(1998).
DR EMBL: AF060999; G3639041; -;
SQ SEQUENCE 391 AA; 41284 MW; 3DD4B2F9 CRC32;

Query Match 10.1%; Score 107.5; DB 2; Length 391;
Best Local Similarity 24.5%; Pred. No. 0.039;
Matches 47; Conservative 31; Mismatches 79; Indels 35; Gaps 7;

OY 22 GTGCGSQNGEGTPTWOSLEPOSSDALEAMENGERSRRRAVLTQKKOHSVHLVPINAT 81
DB 221 GPPGP-----OGPPGLQCGSGADKAGTREN-----OPAVVHLQCGGSA 259
OY 82 --SKD-----SDTYEVMWOPALRRGRGLQAQGYVRIQDAGVLLYSQVLFODVFTM 133
DB 260 IOVKNDLSGGVLDMSRTIMNPKVKR--LHPRSGELEVDGTFTYSQVEVYINFT 316
OY 134 GOVVSREGGROETLFRICIRSMPSHPDRAYNSQYSGVPHLQGDILSVIIPRARKNL 193
DB 316 -DEAYEVVVDKPEPLQCTRSIETGKTN-YNTCYTAGVCLLKAROKIAVKVHADISIMM 373
OY 194 SPHGTFLEFVKL 205
DB 374 SKHTTFFGAIRL 385

RESULT 3
OY 043647 PRELIMINARY; PRT: 157 AA.
AC 043647;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DE 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR ALPHA (FRAGMENT).
GN TNFA.
OS HOMO SAPIENS (HUMAN).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; PRIMATES;
OC CATARRHINI; HOMINIDAE; HOMO.
RN [1]
RP SEQUENCE FROM N.A.
RC JANG J.S., KIM B.E.;
RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF043342; G2905634; -;
DR PROSITE: PS00251; TNF.1; 1.
FT NON_TER 1
SQ SEQUENCE 157 AA; 17380 MW; FD78F0A6 CRC32;

Query Match 8.5%; Score 90.5; DB 2; Length 157;
Best Local Similarity 28.3%; Pred. No. 0.42;
Matches 39; Conservative 20; Mismatches 50; Indels 29; Gaps 7;

OY 78 INATSKDSD--VTEVMWOPAL-----RRGRGLQAQGYVR-----IODAGVLLYS 122
DB 1 VSSSRTPSDKVAHVAVNAPQAEGLQWLMRRANMLLNGVLRDNQVLPVSEGLYLIS 60
OY 123 QVLFPD-----VTFMGV-VSRREGQROETLFR--CIRSMPSHPDRA--YNSQYS 168
DB 61 OVLFRGGGCPSTHVLTLTISRIVASYQTKVLLSLAIKSPCORETPRGAARPWEPYIL 120

OY 169 AGVPHLHOGDILSVIIPR 186
DB 121 GGVFPLEKGDRLSAFNR 138

RESULT 4
OY 054693 PRELIMINARY; PRT: 377 AA.
AC 054693;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE EDA PROTEIN HOMOLOG.
GN TABBY.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIROGNATHI; MORIDAE; MORINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE: 97449184.
RA FERGUSON B.M., BROCKDORF N., FORMSTONE E., NGUYEN T., KRONMILLER J.E.,
RA ZONANA J.;
RT "Cloning of Tabby, the murine homolog of the human EDA gene: evidence
RT for a membrane-associated protein with a short collagenous domain.";
RL HUM. MOL. GENET. 6:1589-1594(1997).
RN [2]
RP SEQUENCE OF 1-132 FROM N.A.
RC STRAIN-129;
RA BROCKDORF N., ZONANA J., FORMSTONE E., FERGUSON B.;
RL SUBMITTED (MAY-1997) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF004435; G2660700; -;
DR EMBL: AF004434; G2660698; -;
SQ SEQUENCE 377 AA; 40002 MW; AEA7ED8 CRC32;

Query Match 8.4%; Score 89.5; DB 10; Length 377;
Best Local Similarity 25.6%; Pred. No. 1.5;
Matches 33; Conservative 26; Mismatches 49; Indels 21; Gaps 5;

OY 94 OPALRRGRGLQAQGYVRID--AGVLLYSQVLFODVFTM--GOV----- 137
DB 247 OPVVH--LOGGSAIGVKNDLSGGVLDMSRTIMNPKVKFLHPRSGELEVYINFTDF 303
OY 137 VSREGGROETLFRICIRSMPSHPDRAYNSQYSGVPHLQGDILSVIIPRARKNLSPH 196
DB 304 ASYEYVVDKPEPLQCTRSIETGKTN-YNTCYTAGVCLLKAROKIAVKVHADISIMSKH 362
OY 197 GTFLEFVKL 205
DB 363 TTFGAIRL 371

RESULT 5
OY 043143 PRELIMINARY; PRT: 583 AA.
AC 043143;
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE PECTINESTERASE (EC 3.1.1.11) (PECTIN METHYLESTERASE) (PE).
GN PMEUI.
OS LYCOPERSTON ESCULENTUM (TOMATO).
OC EUKARYOTA; VIRIDIPLANTAE; CHAROPHYTA/EMBRIOPHYTA GROUP; EMBRIOPHYTA;
OC TRACHEOPHYTES; EUPHYLLOPHYTES; SPERMATOPHYTA; MAGNOLIOPHYTA;
OC EUDICOTYLEDONS; ASTERIDAE; SOLANACEAE; SOLANACEAE; SOLANUM.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RIO GRANDE;
RA GARFE J., TIZNADO M.E., HANDA A.K.;
RL PLANT PHYSIOL. 110:1436-1436(1996).
DR EMBL: U49330; G1222552; -;
DR PROSITE: PS00503; PECTINESTERASE_2; 1.


```

DB 137 QVLFKGGGCPNHYLTHHTTSRIVSYOTKYVNLSTAIRKSPCQRETPGABAKWPYETYL 196
QY 169 AGVPHLHOGDILSYIPRAKAKLNLSPHGT 199
DB 197 GGVFOLEKGRDRLS-----AEINLPDYLD 220

RESULT 9
Q94537
ID 094537 PRELIMINARY; PRT; 1375 AA.
AC 094537;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE FRAZZLED.
GN FRAZZLED.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYROIDERA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE: 97015076.
RA KOLODZIEJ P.A., TIMPE L., MITCHELL K.J., GOODMAN C.S., FRIED S.,
  "frazzled encodes a Drosophila member of the DCC immunoglobulin
  subfamily and is required for CNS and motor axon guidance.";
RT CELL 87:197-204(1996)
RL EMBL: U71001; G1621117; -.
DR FLYBASE; FBgn0011592; fra.
DR PFAM; PF00041; fn3; 6.
DR PFAM; PF00047; 1g; 3.
SQ SEQUENCE 1375 AA; 151692 MW; 43806DBC CRC32;

```

```

Query Match 8.1%; Score 85.5; DB 3; Length 1375;
Best Local Similarity 22.5%; Pred. No. 17;
Matches 41; Conservative 13; Mismatches 47; Indels 81; Gaps 7;

QY 6 QOTELQSL-----REVSRLOGTG-----GPSQNGEGYPMOSLPE 40
DB 516 QQVNIOSLPGRTYQFVEANTNGSGASSAPLEVSTQPEVNINAGPPRNEGY-----AR 570
QY 41 QSSDALEAME-----NGERSRKR-----RAVLQKKOKKHSVYLHV 76
DB 571 SHKEIYKWEPTVTNGEILKRYVYSENDSGADLYHDSTALEVLTLPHTDYIVSV 630
QY 77 PINATSKDSD-----DYTE-----VMMQPALRGRGLQAQGYG 109
DB 631 PFNRNGMGDSASAIRVKTFSSTPEPPNNVTLEVYSSSSITVHWPPEAEDRNGQITGYK 690
QY 110 VR 111
DB 691 IR 692

RESULT 10
Q94538
ID 094538 PRELIMINARY; PRT; 1526 AA.
AC 094538;
DT 01-FEB-1997 (TREMBLREL. 02, CREATED)
DT 01-FEB-1997 (TREMBLREL. 02, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE FRAZZLED.
GN FRAZZLED.
OS DROSOPHILA MELANOGASTER (FRUIT FLY).
OC EUKARYOTA; METAZOA; ARTHROPODA; TRACHEATA; HEXAPODA; INSECTA;
OC PTERYGOTA; DIPTERA; BRACHYCERA; MUSCOMORPHA; EPHYROIDERA;
OC DROSOPHILIDAE; DROSOPHILA.
RN [1]
RP SEQUENCE FROM N.A.
MEDLINE: 97015076.

```

```

RA KOLODZIEJ P.A., TIMPE L., MITCHELL K.J., GOODMAN C.S., FRIED S.,
  "frazzled encodes a Drosophila member of the DCC immunoglobulin
  subfamily and is required for CNS and motor axon guidance.";
RT CELL 87:197-204(1996)
RL EMBL: U71002; G1621117; -.
DR FLYBASE; FBgn0011592; fra.
DR PFAM; PF00041; fn3; 6.
DR PFAM; PF00047; 1g; 3.
SQ SEQUENCE 1526 AA; 168787 MW; F17B1EC9 CRC32;

```

```

Query Match 8.1%; Score 85.5; DB 3; Length 1526;
Best Local Similarity 22.5%; Pred. No. 20;
Matches 41; Conservative 13; Mismatches 47; Indels 81; Gaps 7;

QY 6 QOTELQSL-----REVSRLOGTG-----GPSQNGEGYPMOSLPE 40
DB 667 QQVNIOSLPGRTYQFVEANTNGSGASSAPLEVSTQPEVNINAGPPRNEGY-----AR 721
QY 41 QSSDALEAME-----NGERSRKR-----RAVLQKKOKKHSVYLHV 76
DB 722 SHKEIYKWEPTVTNGEILKRYVYSENDSGADLYHDSTALEVLTLPHTDYIVSV 781
QY 77 PINATSKDSD-----DYTE-----VMMQPALRGRGLQAQGYG 109
DB 782 PFNRNGMGDSASAIRVKTFSSTPEPPNNVTLEVYSSSSITVHWPPEAEDRNGQITGYK 841
QY 110 VR 111
DB 842 IR 843

RESULT 11
Q54910
ID 054910 PRELIMINARY; PRT; 364 AA.
AC 054910;
DT 01-JUN-1998 (TREMBLREL. 06, CREATED)
DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE NUCLEAR FACTOR OF KAPPA LIGHT POLYPEPTIDE PROTEIN ENHANCER IN B-CELLS
  INHIBITOR, EPSILON (IKB EPSILON).
GN NFKBIE.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIUROGNATHI; MORIDAE; MORINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RC MEDLINE: 98070758.
RA SIMONIDIS S., LIANG S., CHEN G., THANOS D.,
  "Cloning and functional characterization of mouse IkappaBepsilon.";
RL PROC. NATL. ACAD. SCI. U.S.A. 94:14372-14377(1997).
DR EMBL; AF030896; G2739158; -.
DR MGI; MGI:1194908; NFKBIE.
SQ SEQUENCE 364 AA; 39327 MW; EC4923B7 CRC32;

```

```

Query Match 8.1%; Score 85.5; DB 10; Length 364;
Best Local Similarity 23.1%; Pred. No. 3; 3;
Matches 46; Conservative 30; Mismatches 82; Indels 41; Gaps 9;

QY 10 LOSLR--EVSRLQGTGPSQNGEGYPMOSLPPQSSDALEAMENGRSRRRAVLTOKOK 67
DB 23 LRSLSRSLPEPTAAGSGS--SQSGCPQFWRHAPETHKEPEKEDADGERADSTYA----- 75
QY 68 KQSVLHVLI--NATSKDSDVTEYMMQPALRGR-----RGLQAQGYVRIQDAGYVYL 121
DB 75 --SSLTSEFPLERPEKADS--PVPVGSPLPAGVLSFOQLEALY----ISDGGDTLH 128
QY 122 SOYLFQDVTFTMGQVVSREGQROETLFRICRSPMSPHPRAYNSQYSAGV--PHLHOGD 178
DB 129 LAVIHE-----APSVLFCCLATLPQEVLDIQNNVLTOTALHVAVHLDQDP 172

```


OY 179 ILSVIIPARAKLNLSPHG 197
Db 173 VVRAVLKGSASRILODQH 191

RESULT 12

093529 PRELIMINARY: PRT: 319 AA.
ID 093529
AC 093529
DT 01-NOV-1998 (TREMBLREL. 08, CREATED)
DT 01-NOV-1998 (TREMBLREL. 08, LAST SEQUENCE UPDATE)
DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE FORKHEAD-DOMAIN-CONTAINING PROTEIN 5.
CN FKH-5.
OS XENOPUS LAEVIS (AFRICAN CLAMED FROG).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; AMPHIBIA; BATRACHIA; ANURA;
OC MESOBATRACHIA; PIPOIDEA; PIPOIDAE; XENOPODINAE; XENOPUS.
RN [1]
RP SEQUENCE FROM N.A.
RA GANSE J.T., KUO J., PATEL M., SIVE H.;
RT "Patterning of the xenopus ectoderm during blastula and gastrula stages."
RL SUBMITTED (MAY-1998) TO EMBL/GENBANK/DBJ DATA BANKS.
DR EMBL: AF064810; G3695057; -;
DR PROSITE: PS00657; FORK_HEAD_1; 1.
DR PROSITE: PS00658; FORK_HEAD_2; 1.
SQ SEQUENCE 319 AA; 34954 MW; B2A80638 CRC32;

Query Match 8.1%; Score 85.5; DB 12; Length 319;
Best Local Similarity 23.4%; Pred. No. 2.8;
Matches 43; Conservative 23; Mismatches 85; Indels 33; Gaps 8;

OY 26 PSQNGEGYPMOSLPEQSSDALEAWNGERSKRRRAVLTKQKKQSHVHLVPIINTSKD 85
Db 78 PDQPKGSGFALHPRCG---DMFENGSLRRKRRKFKYKWD-----HLAPSKA----- 123
OY 86 SDVTEVMQPALRRGRGLQAGGYRIODAGYLL-YSQVLEFDYFTMGQVVSREGGR 144
Db 123 SDAQVILQQAQKLRLSALAAS--GTHLPMSYTNLGVSPTSFKKPFALIENTIAIEYKMP 180
OY 145 QETLFRCIRSPSHPRAY-----NSQYSAGVPHLHOGDILSVIIP--RAKAKLNL 193
Db 181 GGLAESTMQMPMA---AYPLHMLTGVGSGTGMPHMTSSMDSTPIISMANSYSV 236
OY 194 SPHG 197
Db 237 SAYG 240

RESULT 13

089250 PRELIMINARY: PRT: 754 AA.
ID 089250
AC 089250
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1996 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE 84 KDA READTHROUGH PROTEIN.
OS SOIL-BORNE WHEAT MOSAIC VIRUS.
OC VIRUSES; SSRNA POSITIVE-STRAND VIRUSES, NO DNA STAGE; FURIOVIRUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-US-NEBRASKA, 1988 WILD-TYPE;
RX MEDLINE: 93303914.
RA SHIRARO Y., WILSON M.A.;
RT "Complete nucleotide sequence and organization of the bipartite RNA genome of soil-borne wheat mosaic virus."
RL VIROLOGY 195:16-32(1993).
DR EMBL: L07938; G295512; -;
SQ SEQUENCE 754 AA; 83736 MW; B41CB937 CRC32;

Query Match 7.8%; Score 83; DB 11; Length 754;
Best Local Similarity 26.1%; Pred. No. 14;
Matches 46; Conservative 24; Mismatches 68; Indels 38; Gaps 9;

OY 13 LRREYSRLQGTGSPQNGEGYPMQSLPEQS-----SDALEAWNG-ERSKRRAY 61
Db 580 LLRALSMLEPSYGGGGGNGCG---GSLPPQALELPDRAFTFERLAALONGIDLSFENNEV 636
OY 62 LTQKKQSHVHLVPIINATSKDSDVTEVMQPALRRGLQAGGYVRI----QDAG 116
Db 637 FTPEELKT-EIRHVVOAYADSS---YHVADEPYLR-----GVGVTVVTGLPEAVG 682
OY 117 -VYLLSYVLEFDYFTMGQVVSREGGRQETLFCISMSHPDRANSQYSAGV 171
Db 683 EITTTSESVSDVTSISLGV---PGRARRRAVSXNTBTPSTGQSQSYNSKSPRV 734

RESULT 14

035853 PRELIMINARY: PRT: 232 AA.
ID 035853
AC 035853
DT 01-JAN-1998 (TREMBLREL. 05, CREATED)
DT 01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE TUMOR NECROSIS FACTOR ALPHA.
CN TNFA.
OS MUS MUSCULUS (MOUSE).
OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIROGNATHI; MURIDAE; MURINAE; MUS.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-A/J;
RX MEDLINE: 97246744.
RA IRAQI F., TEALE A.;
RT "Cloning and sequencing of the tnfa genes of three inbred mouse strains."
RL IMMUNOGENETICS 45:459-461(1997).
DR EMBL: U68414; G2304957; -;
DR PROSITE: PS00251; TNF_1; 1.
DR PFM: PF00229; TNF; 1.
SQ SEQUENCE 232 AA; 25513 MW; 9B2B3F06 CRC32;

Query Match 7.7%; Score 82; DB 10; Length 232;
Best Local Similarity 26.3%; Pred. No. 3.9;
Matches 52; Conservative 24; Mismatches 70; Indels 52; Gaps 12;

OY 22 GTGSPQNGEGYPMOSLPEQSSDALEAWNGERSKRRRAVLTKQKKQSHVHLVPIINAT 81
Db 54 GVIGP-QDEKFP-NGLEPLISSMA-----QTLSSQNSSDKPAHAVVANHQV 98
OY 82 SKDSDVTEVMQPALRRGRGLQAGGYVR-----IDAGVYLLSOVLEFD-----VT 130
Db 99 EE-----QLEM--LSQANALLANGMDKDNQVLPADGLYVLSVLEFGKGGCPDYVL 150
OY 131 FTMGQVVSREGGRQETLFCIRRS--MPSHPDRA-----YNSQYSAGVPHLHOGDILS 181
Db 151 LT--HTVSREFAISYQEKVNLISAVKSPCKPTPEGAELKPWEPIYLLGVRQLKGDQLS 208
OY 182 VIIPARAKLNLSPHGT 199
Db 209 -----AEVNLPKYIDF 219

RESULT 15

031277 PRELIMINARY: PRT: 392 AA.
ID 031277
AC 031277
DT 01-NOV-1996 (TREMBLREL. 01, CREATED)
DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)
DE CELL SURFACE ANTIGEN PRECURSOR.
OS RATTUS NORVEGICUS (RAT).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; RODENTIA;
OC SCIROGNATHI; MURIDAE; MURINAE; RATTUS.
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE: 92250119.
RA KIRISITS M.J., KUNZ H.W., CORTESI HASSETT A.L., GILL T.J. III;
RT "Genomic DNA sequence and organization of a TL-like gene in the
RL IMMUNOGENETICS 35:365-377(1992).
DR EMBL: M74822; G205462; -
DR PFAM: PF00047; Ig; 1.
DR PFAM: PF00129; MHC_I; 1.
KW SIGNAL; MHC.
FT SIGNAL 1 30 POTENTIAL.
FT CHAIN 31 392 CELL SURFACE ANTIGEN.
SQ SEQUENCE 392 AA; 44744 MW; 9C2990D8 CRC32;

Query Match 7.6%; Score 81; DB 5; Length 392;
Best Local Similarity 22.0%; Pred. No. 9.2;

Matches 37; Conservative 22; Mismatches 61; Indels 48; Gaps 8;

QY 34 PMQSLEQSSDALLEAWENGRSRKRAVLTQKQKQSHVLTLPINATSKDSDVTEVMW 93
DB 80 PM-----QDEGLEWE-----QVHKVKTQEQSARNLMMLVRFYKSKMDFHT--LQW 127
QY 94 OPALRRGRGLQAQGYGVRIQDAGVLLYSQVLFQDY-----FTMGQVYS 138
DB 128 Q-----QGCDEVG--SDGRLHWIDQLAFDGYVDHPTLNKDLRFWTAWTSTVAQISO 175
QY 139 RSGQGRQETLFCIRSMPSHPDRAYNSQYSAGVPHLH-----QGD 179
DB 176 PELFARLKD--NCSELLQKYPEKEKERELLRSDEPRRAHVTRQPRPEGDV 221

Search completed: May 14, 1999, 21:40:00
Job time: 7994 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 03:30:53 ; Search time 2254.51 Seconds
(without alignments)
1417.049 Million cell updates/sec

Title: US-09-212-270-8

Perfect score: 1 tccggcgccgcggggcagga.....taaaagtgtctatacccta 893

Scoring table:

Searched: 808301 segs, 1788773984 residues

Database :

GenEmbl.*
1: gb_dal.*
2: gb_da2.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl1.*
9: gb_pl2.*
10: gb_pr1.*
11: gb_pr2.*
12: gb_pr3.*
13: gb_ro.*
14: gb_st.*
15: gb_sy.*
16: gb_un.*
17: gb_vl.*
18: gb_vtg.*
19: em_da.*
20: em_fun.*
21: em_hum1.*
22: em_hum2.*
23: em_in.*
24: em_om.*
25: em_or.*
26: em_ov.*
27: em_pat.*
28: em_ph.*
29: em_pl.*
30: em_ro.*
31: em_sy.*
32: em_un.*
33: em_vl.*
34: em_vtg.*
35: em_st.*
36: gb_dal.*
37: gb_da2.*
38: gb_pl1.*
39: gb_pl2.*
40: gb_pr1.*
41: gb_pr2.*
42: gb_pr3.*
43: gb_st.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Query | Score | Match | Length | DB | ID | Description |
|------------|-------|-------|-------|--------|----|----|-------------|
|------------|-------|-------|-------|--------|----|----|-------------|

| | | | | | | |
|----|------|-----|--------|----|-------------|----------------------|
| 1 | 36.8 | 4.1 | 3520 | 3 | DMU23545 | U23545 Drosophila |
| 2 | 36.8 | 4.1 | 9894 | 11 | HS113D11 | Z83837 Human DNA s |
| 3 | 36.8 | 4.1 | 9894 | 41 | HS113D11 | Z83837 Human DNA s |
| 4 | 36.6 | 4.0 | 3339 | 10 | HUM8DC62 | L35661 Homo sapien |
| 5 | 36.6 | 4.0 | 60904 | 12 | AC005612 | AC005612 Homo sapi |
| 6 | 36.6 | 4.0 | 110000 | 18 | HSAT11929_5 | Continuation (6 of |
| 7 | 36.6 | 4.0 | 3339 | 40 | HUM8DC62 | L35661 Homo sapien |
| 8 | 36.6 | 4.0 | 60904 | 42 | AC005612 | AC005612 Homo sapi |
| 9 | 35.6 | 4.0 | 140977 | 11 | AC002536 | AC002536 Human Chr |
| 10 | 35.6 | 4.0 | 140977 | 41 | AC002536 | AC002536 Human Chr |
| 11 | 35.4 | 4.0 | 2315 | 1 | BEU75371 | U75371 Bacteroides |
| 12 | 35.4 | 4.0 | 2315 | 36 | BEU75371 | U75371 Bacteroides |
| 13 | 35.3 | 3.9 | 1912 | 13 | MMRARG | X15848 Mouse mRNA |
| 14 | 34.8 | 3.9 | 40127 | 11 | HSF0811 | Z97184 Human DNA s |
| 15 | 34.8 | 3.9 | 39872 | 11 | HSICB2046 | Z97183 Human DNA s |
| 16 | 34.8 | 3.9 | 40127 | 41 | HSF0811 | Z97184 Human DNA s |
| 17 | 34.8 | 3.9 | 39872 | 41 | HSICB2046 | Z97183 Human DNA s |
| 18 | 34.4 | 3.9 | 4018 | 10 | HUMRPO04 | M25704 Human thryo |
| 19 | 34.4 | 3.9 | 115835 | 12 | HS28C20 | AL031119 Human thryo |
| 20 | 34.4 | 3.9 | 4018 | 40 | HUMRPO04 | M25704 Human thryo |
| 21 | 34.4 | 3.9 | 115835 | 42 | HS28C20 | AL031119 Human thryo |
| 22 | 34.4 | 3.9 | 115835 | 42 | HS28C20 | AL031119 Human thryo |
| 23 | 34.4 | 3.8 | 11335 | 1 | AE000082 | AE000082 Rhizobium |
| 24 | 34.4 | 3.8 | 3233 | 2 | TTARGOPE | Y10525 T.thermophi |
| 25 | 34.4 | 3.8 | 6824 | 2 | TTARGOPE | Y10525 T.thermophi |
| 26 | 34.4 | 3.8 | 11335 | 36 | AE000082 | AE000082 Rhizobium |
| 27 | 34.4 | 3.8 | 3233 | 36 | TTARGOPE | Y10525 T.thermophi |
| 28 | 34.4 | 3.8 | 6824 | 37 | TTARGOPE | Y10525 T.thermophi |
| 29 | 33.8 | 3.8 | 3013 | 4 | RABTLCCEPH | L13199 Rabbit tele |
| 30 | 33.8 | 3.8 | 3013 | 27 | E09934 | E09934 CDNA encod1 |
| 31 | 33.6 | 3.8 | 3851 | 12 | HSITGBF07 | U66535 Human beta4 |
| 32 | 33.6 | 3.8 | 3851 | 42 | HSITGBF07 | U66535 Human beta4 |
| 33 | 33.4 | 3.7 | 201577 | 18 | AC005831 | AC005831 *** SEQUE |
| 34 | 33.4 | 3.7 | 1376 | 6 | I01659 | I01659 Sequence 8 |
| 35 | 33.3 | 3.7 | 2810 | 1 | YSJLEU2B | M37309 Y. lipolytic |
| 36 | 33.3 | 3.7 | 44762 | 11 | AC002325 | AC002325 Homo sapi |
| 37 | 33.3 | 3.7 | 162485 | 18 | AC004840 | AC004840 *** SEQUE |
| 38 | 33.3 | 3.7 | 2810 | 18 | YSJLEU2B | M37309 Y. lipolytic |
| 39 | 33.3 | 3.7 | 44762 | 41 | AC002325 | AC002325 Homo sapi |
| 40 | 32.8 | 3.7 | 189370 | 1 | AF010496 | AF010496 Rhodobact |
| 41 | 32.8 | 3.7 | 4809 | 1 | AF017138 | AF017138 Rhodobact |
| 42 | 32.8 | 3.7 | 2070 | 1 | RHBSRA | M73823 Bradyrhizob |
| 43 | 32.8 | 3.7 | 4998 | 11 | AF001042 | AF001042 Homo sapi |
| 44 | 32.8 | 3.7 | 211072 | 18 | HS1108R3 | AL033525 Human DNA |
| 45 | 32.8 | 3.7 | 189370 | 36 | AF010496 | AF010496 Rhodobact |

ALIGNMENTS

| RESULT | 1 | | | | | |
|------------|---|---------|------|-----|-------------|--|
| DMU23545 | | 3520 bp | mRNA | INV | 19-JUL-1995 | |
| LOCUS | | | | | | |
| DEFINITION | | | | | | |
| U23545 | | | | | | |
| ACCESSION | | | | | | |
| NID | | | | | | |
| KEYWORDS | | | | | | |
| SOURCE | | | | | | |
| ORGANISM | | | | | | |

fruit fly.
Drosophila melanogaster
Eukaryotae; mitochondrial eukaryotes; Metazoa; Arthropoda;
Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

Submitted (28-MAR-1995) Scott W. Chouinard, Genetics, University of Wisconsin-Madison, 445 Henry Mall, Madison, WI 53706, USA

| | |
|------------|---|
| RESULT | 3 |
| LOCUS | H5113D11 |
| DEFINITION | H5113D11 9894 bp DNA PRI 21-AUG-1998 |
| ACCESSION | Human DNA sequence from fosmid 113D11 on chromosome 22 contains ESTs, Cpg island, complete sequence. |
| NID | 283837 |
| KEYWORDS | g2578046 |
| SOURCE | HTG: Cpg Island. |
| ORGANISM | human. |
| REFERENCE | Homo sapiens |
| AUTHORS | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; |
| TITLE | Primates; Catarrhini; Hominoidea; Homo. |
| JOURNAL | 1 (bases 1 to 9894) |
| COMMENT | Burgess,J Direct Submission Submitted (15-OCT-1997) E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Oct 31, 1997 this sequence version replaced gi:1772930. IMPORTANT: This sequence is not the entire insert of clone 113D11. It may be shorter because we only sequence overlapping sections one, or longer because we arrange for a small overlap between neighbouring submissions. This sequence has been finished according to sequence map criteria as follows. An attempt is made to resolve all sequencing problems, such as compressions and repeats, but not necessarily within known annotated human repeat sequence elements (e.g. Alu). Where the sequence is ambiguous, there is an annotation using the "unsure" feature key. The true left end of clone 113D11 is at 1 in this sequence. The true left end of clone N104C4 is at 9791. This sequence was generated from part of bacterial clone conlgs of human chromosome 22, constructed by the Sanger Centre Chromosome 22 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Ctr22 . Location/Qualifiers 1..9894 /organism="Homo sapiens" /db_xref="taxon:9606" /chromosome="22" /clone="113D11" 74..139 /note="MLRD repeat: matches 327..393 of consensus" 808..2279 /note="Putative Cpg Island" prim_transcript 833..1251 /note="match: 5' EST AA173131 clone 611044" prim_transcript <1279..1570 /note="match: EST D21359" prim_transcript 2069..>2815 /note="match: multiple ESTs: match: H40796 H88498" repeat_region 4787..4883 /note="MIR repeat: matches 44..146 of consensus" 5509..5559 /note="3 copies of 17 mer 88 & conserved" 5625..5745 /note="C rich, A poor region" 5746..5805 /note="2 copies of 30 mer 93 & conserved" 5815..5898 /note="3 copies of 28 mer 87 & conserved" 6281..6292 /note="Single clone region." 8013..8355 /note="G, A rich region" 8179..8291 /note="Single clone region." 8325..8332 /note="False join. Repeat sequence of unknown length." 8333..8377 /note="Single clone region." BASE COUNT 1937 a 2837 c 3076 g 2044 t ORIGIN |

| | | | | |
|-----------------------|---|--|-----------|---------------|
| Query Match | 4.1%: | Score 36.8: | DB 41: | Length 9894: |
| Best Local Similarity | 61.5%: | Pred. No. 3.6: | | |
| Matches 59: | Conservative 0: | Mismatches 37: | Indels 0: | Gaps 0: |
| OY | 502 | accagaagtgtaacctgtgtaagatctctctccacgcctccagccctcggagacagcg | 561 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db | 7353 | ATCGCAGAGGAGACCTCATCAACACCCGCTACACAGGGGCGAGGACTGAGAGGTGGCG | 7412 | |
| OY | 562 | gtggtgcgcgtcatctcgtcaagtcgtcttcggccagctg | 597 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db | 7413 | TGAGATGGGCTCAGCTCCAACCCACTGGCTCGGCG | 7448 | |
| RESULT 4 | | | | |
| HUM8DC6/c | | | | |
| LOCUS | HUM8DC62 | 3339 bp | DNA | 22-AUG-1994 |
| DEFINITION | Homo sapiens (subclone H8_4_b9 from P1 35 H5 C8) | | | DNA sequence. |
| ACCESSION | L35661 | | | |
| NID | 9532026 | | | |
| KEYWORDS | MX1 region. | | | |
| SOURCE | Homo sapiens (Library: Subclones in pSP72 from P1 clone 35 H5 C8 (H8)) | | | DNA. |
| ORGANISM | Homo sapiens | | | |
| | Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Euteria; Primates; Catarrhini; Homiidae; Homo. 1 (bases 1 to 3339) | | | |
| REFERENCE | Martin,C.H., Bondoc,M.M., Chang,A., Cloutier,T., Davis,C.A., Ericsson,C.L., Jaklevic,M.A., Kim,R.J., Lee,M.T., Li,M., Mayeda,C.A., Steiert-El Khelr,A. and Palazzolo,M.J. | | | |
| AUTHORS | Sequencing of the MX1 region on human chromosome 21 unpublished (1994) | | | |
| TITLE | Sequence submitted by: | | | |
| JOURNAL | Human Genome Center and | | | |
| COMMENT | Human Genome Center and | | | |
| | Drosophila Genome Center | | | |
| | Lawrence Berkeley Laboratory | | | |
| | Berkeley, CA 94720 | | | |
| | e-mail: seqgenome.lbl.gov | | | |
| | This subclone overlaps H8_6_e2 and H8_3_d9. | | | |
| FEATURES | Location/Qualifiers | | | |
| source | 1..3339 | | | |
| | /organism="Homo sapiens" | | | |
| | /db_xref="taxon:9606" | | | |
| | /tissue_lib="Subclones in pSP72 from P1 clone 35 H5 C8 (H8)" | | | |
| BASE COUNT | 749 a 969 c 875 g 746 t | | | |
| ORIGIN | | | | |
| Query Match | 4.0%: | Score 36: | DB 10: | Length 3339: |
| Best Local Similarity | 49.0%: | Pred. No. 5.3: | | |
| Matches 96: | Conservative 0: | Mismatches 100: | Indels 0: | Gaps 0: |
| OY | 228 | gtccagaagatgtgtgcttcggtctatgagggaagatgcacagtgtgtgacctcagggcgca | 287 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db | 332 | GTACATATATTTAAAGGAGCTATAGGGTTAAAGAACCCACAGTGGAGGCTCTGGAGGCCAGC | 333 | |
| OY | 288 | ccggttcagaagaagacttgggtttccagaagtgttaagccatgtgtgagactgtgcgtgt | 347 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db | 332 | CCACTGACAGCAGCCACTCCAGGAGAGTCCAAAGTCCCGCTCTAGTGTGCTGGGTGAGGGA | 273 | |
| OY | 348 | gaacggtcttcaggaggcgcaactgttcacacacacagtgatgtctgtcgtggagactgtct | 407 | |
| | | | | |
| | | | | |
| | | | | |
| | | | | |
| Db | 272 | AATCTGTTCTCCAGGAGCCTGTCTCTGGCTGCCCACTCTCCAAAGTCCAGGAATTAAGCT | 213 | |
| OY | 408 | gccagagatttaccgg 423 | | |
| | | | | |
| | | | | |
| Db | 212 | TTCAGAAATTCACATG 197 | | |
| RESULT 5 | | | | |
| NC005612 | | | | |

| | | | | | |
|---------------|---|----------|-----|-----|-------------|
| LOCUS | AC005612 | 60904 bp | DNA | PRI | 04-SEP-1998 |
| DEFINITION | Homo sapiens chromosome 21, p1 clone LBI#8 (DBNL H8), complete sequence. | | | | |
| ACCESSION | AC005612 | | | | |
| NID | 93540153 | | | | |
| KEYWORDS | HTG. | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 60904) Kimberly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M., Kader K., Mignel T., Miller C., Pitluck S., Pollard M., Rojesti H., Subramanian S. and Martin C.H. | | | | |
| TITLE | Sequencing of human chromosome 21 | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 2 (bases 1 to 60904) | | | | |
| AUTHORS | Ricke D.O. | | | | |
| TITLE | Large Scale Sequence Analysis and Annotation with the Sequence Comparison Analysis (SCAN) System | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 3 (bases 1 to 60904) | | | | |
| AUTHORS | Kimberly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M., Davis C.A., Kader K., Mignel T., Pitluck S., Pollard M., Rojesti H., Subramanian S. and Martin C.H. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (04-SEP-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A. | | | | |
| COMMENT | Sequence submitted by: DOE Joint Genome Institute. | | | | |
| FEATURES | Location/Qualifiers | | | | |
| Source | 1..60904 /organism="Homo sapiens" /db_xref="taxon:9606" /map="21" /clone="p1 LBI#8" /chromosome="21" /note="LBNL H8" | | | | |
| repeat_region | 1..675 /rpt_family="MER25" | | | | |
| repeat_region | 423..565 /rpt_family="L1" | | | | |
| repeat_region | 815..4661 /rpt_family="L1" | | | | |
| repeat_region | 4917..4940 /note="(ACT)8" | | | | |
| repeat_region | /rpt_type=tandem /rpt_unit=ACT complement(5525..5890) /rpt_family="THE1" | | | | |
| repeat_region | complement(6688..6819) /rpt_family="Alu" | | | | |
| repeat_region | 7441..7741 /rpt_family="Alu" | | | | |
| repeat_region | 7724..7743 /note="(A)20" | | | | |
| repeat_region | /rpt_type=tandem /rpt_unit=A complement(8063..8123) /rpt_family="MER5" | | | | |
| repeat_region | complement(8811..9778) /rpt_family="LTR5" | | | | |
| repeat_region | complement(9970..10409) /rpt_family="MER41" | | | | |
| repeat_region | complement(11266..11344) /rpt_family="LTR12" | | | | |
| repeat_region | 12169..12204 /note="(GT)18" | | | | |
| repeat_region | /rpt_type=tandem /rpt_unit=GT complement(13074..13520) /rpt_family="LTR7" | | | | |
| repeat_region | complement(13974..14029) /rpt_family="MER5" | | | | |
| repeat_region | 15399..15473 /rpt_family="L1" | | | | |
| misc_feature | 15697..15808 /note="GRAIL 2 excellent exon, frame 0" | | | | |
| repeat_region | complement(15858..16136) /rpt_family="Alu" | | | | |
| misc_feature | 16565..16698 /note="GRAIL 2 excellent exon, frame 1" | | | | |
| repeat_region | 17631..17863 /rpt_family="Alu" | | | | |
| repeat_region | complement(18991..19253) /rpt_family="Alu" | | | | |
| repeat_region | complement(19424..19715) /rpt_family="Alu" | | | | |
| misc_feature | 20241..20286 /note="GRAIL 2 excellent exon, frame 0" | | | | |
| repeat_region | 21081..21263 /rpt_family="MER5" | | | | |
| repeat_region | 21861..22013 /rpt_family="MER20" | | | | |
| repeat_region | complement(23391..23728) /rpt_family="THE1" | | | | |
| misc_feature | 23779..23994 /note="GRAIL 2 excellent exon, frame 0" | | | | |
| repeat_region | complement(24012..24087) /rpt_family="MUT1" | | | | |
| misc_feature | 24957..25094 /note="GRAIL 2 excellent exon, frame 1" | | | | |
| repeat_region | 25319..25385 /rpt_family="Alu" | | | | |
| misc_feature | 27635..27825 /note="GRAIL 2 excellent exon, frame 0" | | | | |
| repeat_region | 28141..28430 /rpt_family="Alu" | | | | |
| misc_feature | 28834..28966 /note="GRAIL 2 excellent exon, frame 0" | | | | |
| misc_feature | 29686..29929 /note="GRAIL 2 excellent exon, frame 2" | | | | |
| repeat_region | 29977..30000 /note="(GT)12" | | | | |
| repeat_region | /rpt_type=tandem /rpt_unit=GT 31701..31844 /note="GRAIL 2 excellent exon, frame 2" | | | | |
| misc_feature | complement(31932..32057) /rpt_family="MER5" | | | | |
| repeat_region | 32297..32320 /note="(T)24" | | | | |
| repeat_region | /rpt_type=tandem /rpt_unit=T complement(32303..32508) /rpt_family="Alu" | | | | |
| misc_feature | 33424..33584 /note="GRAIL 2 excellent exon, frame 0" | | | | |
| misc_feature | 33976..34117 /note="GRAIL 2 excellent exon, frame 0" | | | | |
| repeat_region | complement(34362..34538) /rpt_family="T199er1" | | | | |
| repeat_region | complement(34577..34696) /rpt_family="T199er1" | | | | |
| repeat_region | complement(35034..35111) /rpt_family="MUT1" | | | | |
| repeat_region | complement(35713..36022) /rpt_family="Alu" | | | | |
| repeat_region | 36501..36803 /rpt_family="Alu" | | | | |
| repeat_region | 36909..36950 /note="(GT)21" | | | | |
| repeat_region | /rpt_type=tandem /rpt_unit=GT 37127..37266 | | | | |
| misc_feature | | | | | |

| | |
|---------------|--|
| repeat_region | complement(13974..14029) /rpe_family="MER5" 15399..15473 /rpe_family="I1" 15697..15808 /note="GRAIL 2 excellent exon, frame 0" |
| repeat_region | complement(15858..16136) /rpe_family="Alu" 16565..16698 /note="GRAIL 2 excellent exon, frame 1" 17631..17863 /rpe_family="Alu" complement(16891..19253) /rpe_family="Alu" complement(19424..19715) /rpe_family="Alu" 20241..20286 /note="GRAIL 2 excellent exon, frame 0" |
| repeat_region | 21081..21263 /rpe_family="MER5" 21861..22013 /rpe_family="MER20" complement(23391..23728) /rpe_family="THE1" 23779..23994 /note="GRAIL 2 excellent exon, frame 0" |
| repeat_region | complement(24012..24087) /rpe_family="MLT1" 24957..25094 /note="GRAIL 2 excellent exon, frame 1" 25319..25585 /rpe_family="Alu" 27635..27825 /note="GRAIL 2 excellent exon, frame 0" |
| repeat_region | 28141..28430 /rpe_family="Alu" 28784..28966 /note="GRAIL 2 excellent exon, frame 0" 29686..29929 /note="GRAIL 2 excellent exon, frame 2" 29977..30000 /note="(GT)12" /rpe_type=tandem /rpe_unit=GT 31701..31844 /note="GRAIL 2 excellent exon, frame 2" |
| repeat_region | complement(31932..32057) /rpe_family="MER5" 32297..32320 /note="(T)24" /rpe_type=tandem /rpe_unit=T complement(32303..32508) /rpe_family="Alu" 33424..33584 /note="GRAIL 2 excellent exon, frame 0" 33976..34117 /note="GRAIL 2 excellent exon, frame 0" |
| repeat_region | complement(34362..34538) /rpe_family="T199ger1" complement(34577..34696) /rpe_family="T199ger1" complement(35034..35111) /rpe_family="MLT1" complement(35713..36022) /rpe_family="Alu" 36501..36803 /rpe_family="Alu" 36909..36950 /note="(GT)21" /rpe_type=tandem /rpe_unit=GT 37137..37266 |
| misc_feature | |

```

repeat_region /note="GRAIL 2 excellent exon, frame 2"
39543..39822 /rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 1"
40592..40840 /rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
40971..41337 /rpt_family="THE1"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
41617..41717 /rpt_family="MER4"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
42881..43215 /rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
43942..44219 /rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
44724..45031 /rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
46007..46178 /rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 2"
46500..46725 /rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 2"
47500..47781 /rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 2"
49242..49309 /rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 2"
49935..50077 /rpt_family="THE1"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
50080..51967 /rpt_family="MSTAR"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
50244..50543 /rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 1"
51595..51645 /rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
51968..52329 /rpt_family="THE1"
repeat_region /note="GRAIL 2 excellent exon, frame 1"
53080..55378 /rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 0"
56511..56511 /rpt_family="Alu"
misc_feature /note="GRAIL 2 excellent exon, frame 0"
58678..58678 /rpt_family="Alu"
repeat_region /note="GRAIL 2 excellent exon, frame 0"
59339..59470 /rpt_family="MER20"
misc_feature /note="GRAIL 2 excellent exon, frame 1"
59910..59910 /rpt_family="Alu"
BASE COUNT 16756 a 14066 c 14518 g 15564 t
ORIGIN

```

```

Query Match 4.0% Score 36; DB 12; Length 60904;
Best Local Similarity 49.0%; Pred. No. 7.6; Indels 0; Gaps 0;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 228 gtccaaggaatgtgcttcgctatggtggaagatgcacagtgctgcccgcagcgcca 287
DB 16307 GTACATGATTAAAGGAGATGATAGGTAAGACCCACAGTGGAGCTGTGAGGCCAGC 16366
QY 288 ccggttcagaagaagactgggtttccagaagtgaagcactgtgcggactgtgcgtgct 347
DB 16367 CCACGTGACAGCCACTCCAGGAGTCCAGAAATCCGCTAGTGTGCTGGGTGAGGGA 16426
QY 348 gaaccgcttcgaagggcgaactgtctacacacccagtgatgtctgtcgggagactgct 407
DB 16427 AATCGTTCCTCCAGGAGACTGCTCGCTGCCAGCTGCCAAGTCAAGGAAATTAAGCT 16486
QY 408 gccaggaattaccgg 423
DB 16487 TTCAGAAATCTCAGTCG 16502

```

```

RESULT 6
HSAJ11929_5
SEQUENCE split into 7 fragments LOCUS HSAJ11929 Accession AJ011929
WPCOMMENT
Fragment Name Begin End

```

```

HSAJ11929_0 1 110000
HSAJ11929_1 100001 210000
HSAJ11929_2 200001 310000
HSAJ11929_3 300001 410000
HSAJ11929_4 400001 510000
HSAJ11929_5 500001 610000
HSAJ11929_6 600001 614109
Continuation (6 of 7) of HSAJ11929 from base 500001 (AJ011929 Homo sapiens *** SEQUEN

```

```

Query Match 4.0% Score 36; DB 18; Length 110000;
Best Local Similarity 49.0%; Pred. No. 8.1; Indels 0; Gaps 0;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 228 gtccaaggaatgtgcttcgctatggtggaagatgcacagtgctgcccgcagcgcca 287
DB 48708 GTACATGATTAAAGGAGATGATAGGTAAGACCCACAGTGGAGCTGTGAGGCCAGC 48767
QY 288 ccggttcagaagaagactgggtttccagaagtgaagcactgtgcggactgtgcgtgct 347
DB 48768 CCACGTGACAGCCACTCCAGGAGTCCAGAAATCCGCTAGTGTGCTGGGTGAGGGA 48837
QY 348 gaaccgcttcgaagggcgaactgtctacacacccagtgatgtctgtcgggagactgct 407
DB 48838 AATCGTTCCTCCAGGAGACTGCTCGCTGCCAGCTGCCAAGTCAAGGAAATTAAGCT 48887
QY 408 gccaggaattaccgg 423
DB 48888 TTCAGAAATCTCAGTCG 48903

```

```

RESULT 7
LOCUS HUM8DC6Z/c 3339 bp DNA PRI 22-AUG-1994
DEFINITION Homo sapiens (subclone H8_4_b9 from P1 35 H5 C8) DNA sequence.
ACCESSION L35661
KEYWORDS NID
SOURCE MX1 region.
ORGANISM Homo sapiens (library: Subclones in pSP72 from P1 clone 35 H5 C8 (H8)) DNA.
Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Euteria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 3339)
Martin,C.H., Bondoc,M.M., Chiang,A., Cloutier,T., Davis,C.A.,
Mayeda,C.A., Steiert-Ei Kheir,A. and Palazolo,M.J.
Sequencing of the MX1 region on human chromosome 21
Lawrence Berkeley Laboratory
Berkeley, CA 94720
e-mail: seq@genome.lbl.gov
This subclone overlaps H8_6_e2 and H8_3_d9.
FEATURES
location/Qualifiers
1..3339
/organism="Homo sapiens"
/db_xref="taxon:9606"
/tissue_lib="Subclones in pSP72 from P1 clone 35 H5 C8 (H8)"

```

```

BASE COUNT 749 a 969 c 875 g 746 t
ORIGIN

```

```

Query Match 4.0% Score 36; DB 40; Length 3339;
Best Local Similarity 49.0%; Pred. No. 5.3;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;
QY 228 gtccaaggaatgtgcttcgctatggtggaagatgcacagtgctgcccgcagcgcca 287
DB 392 GTACATGATTAAAGGAGATGATAGGTAAGACCCACAGTGGAGCTGTGAGGCCAGC 333

```

[illegible]


```
repeat_region /note="GRAIL 2 excellent exon, frame 0"
complement(34362..34558)
repeat_region /rpt_family="Tiger1"
complement(34577..34696)
repeat_region /rpt_family="Tiger1"
complement(35034..35111)
repeat_region /rpt_family="MLT1"
complement(35713..36022)
repeat_region /rpt_family="Alu"
36501..36803
repeat_region /rpt_family="Alu"
36909..36950
repeat_region /note="(Gr)21"
36953..39822
misc_feature /rpt_type=tandem
37127..37266
misc_feature /rpt_unit=GT
39543..39822
repeat_region /note="GRAIL 2 excellent exon, frame 2"
40592..40840
misc_feature /rpt_family="Alu"
40592..40840
repeat_region /note="GRAIL 2 excellent exon, frame 1"
complement(40971..41337)
repeat_region /rpt_family="THE1"
complement(41617..41717)
repeat_region /rpt_family="MER4"
complement(42881..43215)
repeat_region /rpt_family="Alu"
43942..44219
repeat_region /rpt_family="Alu"
complement(44724..45031)
repeat_region /rpt_family="Alu"
complement(46007..46178)
repeat_region /rpt_family="Alu"
46500..46725
misc_feature /note="GRAIL 2 excellent exon, frame 2"
47500..47781
repeat_region /rpt_family="Alu"
complement(49242..49309)
misc_feature /note="GRAIL 2 excellent exon, frame 2"
49935..50077
repeat_region /rpt_family="THE1"
50080..51967
repeat_region /rpt_family="MSTAR"
complement(50244..50543)
repeat_region /rpt_family="Alu"
complement(51595..51645)
misc_feature /note="GRAIL 2 excellent exon, frame 1"
51968..52329
repeat_region /rpt_family="THE1"
55080..55378
repeat_region /rpt_family="Alu"
complement(56369..56511)
misc_feature /note="GRAIL 2 excellent exon, frame 0"
complement(58621..58678)
misc_feature /note="GRAIL 2 excellent exon, frame 0"
59339..59470
repeat_region /rpt_family="MER20"
complement(59793..59910)
misc_feature /note="GRAIL 2 excellent exon, frame 1"
BASE COUNT 16756 a 14066 c 14518 g 15564 t
ORIGIN
```

```
Db 16367 CCACGTGACAGCACCCTCCAGGAGAGTCACAGAACTCCCGCTCTAGTGTGTGGTGTGTGAGGGA 16426
Qy 348 gaaccgcttcagagggcgaactgtcacacacagatgtctgtcgtcggtgactgct 407
Db 16427 AATCTGTCTCTCCAGGAGGACCTGCTCTGCTGCCACTCCCAAGTCACAGAAATAACT 16486
Qy 408 gccaggaatttaccgg 423
Db 16487 TTCAGAAATCTCACTG 16502

RESULT 9
AC002536 140977 bp DNA 10-DEC-1997
LOCUS Human Chromosome 11 pac pDJ1075f20, complete sequence.
DEFINITION AC002536
ACCESSION g2673898
NID HFG.
KEYWORDS human.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 140977)
AUTHORS Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S.,
Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,
Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P.,
Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
Ward,T. and Wilson,R.
template.
TITLE Unpublished
JOURNAL 2 (bases 1 to 140977)
REFERENCE Evans,G.A., Bradbury,P., Brignac,S., Bumester,R., Burbee,D.,
Davie,J., Davies,C.J., Davis,C., English,C., Fondon,T.,
Franklin,T.L., Garner,H.R., Gordon,M., Gotway,G., Grant,O.,
Hahner,L., Harris,J., Hinson,S., Megarity,C., Narayanaswamy,U.,
Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S.,
Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
Ward,T. and Wilson,R.
Direct Submission
TITLE Submitted (11-SEP-1997) Genome Science and Technology Center,
JOURNAL University of Texas Southwestern Medical Center at Dallas, 5323
Harry Hines Blvd, Dallas, TX 75235-8591, USA
3 (bases 1 to 140977)
REFERENCE Evans,G.A., Athanasiou,M., Basit,M., Bradbury,P., Brignac,S.,
Bumester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R.,
Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Harris,J.,
Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P.,
Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D.,
Ward,T. and Wilson,R.
Direct Submission
TITLE Submitted (10-DEC-1997) Genome Science & Technology Center,
JOURNAL University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA
On Dec 10, 1997 this sequence version replaced gi:2392758.
COMMENT location/Qualifiers
FEATURES 1..140977
source /organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="pDJ1075f20"
/chromosome="11"

BASE COUNT 34355 a 37819 c 37235 g 31568 t
ORIGIN
```

| Query Match | Best Local Similarity | Score | DB | Length |
|--|---|-------------|--------|----------------|
| Matches 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0; | 4.0%; | Score 35.6; | DB 41; | Length 140977; |
| 159 | gcagcgaaggaattcaagagatcgatctggaacggtgtccctctcaacagctgcgacctg | 218 | | |
| 55124 | GGAGGAGCAATTCGACGACCGAGGGGCGAGCAGAGCCCGCTTTGACATCTCGGAGGG | 55183 | | |
| 219 | catgagattgtctcaaggaatctgctctcgatctatgaggagatgacacagttgtgacctg | 278 | | |
| 279 | caggcgcgaccggttcaaggaagacttgggtttccagaagttaagcatgtgcgac | 336 | | |
| 55244 | CCTGCCACGAGAGGCCACGAGAAACCTGCTCTACTGGGGCGCGAGCGGTGCCCCAC | 55243 | | |

| | | | |
|-----------------------|--|--|-----------------|
| Db | 55184 | CCGGAGAGTGATCCCGATGATGATGATTAACCTGGAAACATTTTCACCCCTTCACAGGCCAC | 55243 |
| Qy | 279 | caagcgccagccggttccaagaagactggggtttccagaagtgtlaagccatgtgcggac | 336 |
| Db | 55244 | CCTGCCCAAGAGGCCACGAAACCCCTCCTGTACTGTGGGCCGAGACGCTCCCCAC | 55301 |
| RESULT | 11 | | |
| LOCUS | BFU75371 | 2315 bp | BCU 07-MAY-1997 |
| DEFINITION | Bacteroides fragilis transposon Tn455 transposase (tnpA) gene, | | |
| ACCESSION | complete cds, and transposase (tnpB) gene, partial cds. | | |
| KEYWORDS | U75371 | | |
| SOURCE | 92072415 | | |
| ORGANISM | Bacteroides fragilis. Bacteroides fragilis | | |
| REFERENCE | Eubacteria; Cytophagales; Bacteroidaceae; Bacteroides. | | |
| AUTHORS | 1 (bases 1 to 2315) | | |
| TITLE | tribble,G.D., Parker,A.C. and Smith,C.J. The Bacteroides mobilizable transposon Tn455 integrates by a site-specific recombination mechanism similar to that of the gram-positive bacterial element Tn916 | | |
| JOURNAL | J. Bacteriol. 179 (8), 2731-2739 (1997) | | |
| MEDLINE | 97252504 | | |
| REFERENCE | 2 (bases 1 to 2315) | | |
| AUTHORS | tribble,G.D., Parker,A.C. and Smith,C.J. | | |
| TITLE | Direct Submission | | |
| JOURNAL | Submitted (18-OCT-1996) Microbiology and Immunology, East Carolina University, Moyer Blvd., Greenville, NC 27858-4354, USA | | |
| FEATURES | Location/Qualifiers | | |
| source | 1..2315 | | |
| | /organism="Bacteroides fragilis" | | |
| | /transposon="Tn455" | | |
| | /db_xref="taxon:817" | | |
| stem_loop | 194..224 | | |
| | /function="transposon terminl" | | |
| gene | 569..1504 | | |
| | /gene="tnpA" | | |
| CDs | 569..1504 | | |
| | /gene="tnpA" | | |
| | /codon_start=1 | | |
| | /transl_table=1 | | |
| | /product="transposase" | | |
| | /db_xref="PID:g2072417" | | |
| | /translation="MKATKRCSPGKSFVTRSGMORYCSAPCAQAEAKRVARVOKNNLF KVAPLMEIOHQEYLTESKAAIIMGCRQYIKYVAIGKAKASISRMAFIRADIE OMEGNHYRLPGENISTPFRKSSSSSPARERKEKSEELYDISEEVLSPKVKQ SWLTSKRNIHPICRIAGNYSKSHIDLEFVAVDISITPMLTLEVEELGMP TALAYATRIHKIPIPTREYGRYYSKSHINELRTDLVNDERYYVEQVOOIYGLSSAN ICHTVKKHIEKIKIVGVKNLLRSDVERVAERKK" | | |
| | 1646..2315 | | |
| gene | /gene="tnpB" | | |
| | 1646..2315 | | |
| CDs | /gene="tnpB" | | |
| | /note="Integrase" | | |
| | /codon_start=1 | | |
| | /transl_table=1 | | |
| | /product="transposase" | | |
| | /db_xref="PID:g2072416" | | |
| | /translation="MSKCTVTLTRKKIKNGTOYSLCIDYVPGYRDVNTVRVITREAL GIYIFKAPANDQERDFNARMKKAVIDLRNDRYEIFENNNGFEDTKMGDFLAYFKG LADKNKIMQVHYKHFORVNGKCTFEEDVDICRKMEXILLADPOSIHNNOKHINS AAGWSTFRANLHAYRDKRIKENPFLDRICEIPITRIHLSOEILRIAEIPCEEE VLKRA" | | |
| BASE COUNT | 739 a 533 c 498 g 545 t | | |
| ORIGIN | | | |
| Query Match | 4.0%; Score 35.4; DB 1; Length 2315; | | |
| Best Local Similarity | 49.2%; Pred. No. 7.5; | | |
| Matches | 93; Conservative 0; Mismatches 96; Indels 0; Gaps 0; | | |

```
Qy 692 gaagtagctgttagcatgtcacccaagagttctcaacagacacctgagccttaag 751
|||||
Db 1320 GAACCTGACCTTTGGAACGATGAACGCTACTATACCGTTGACGAGGTGCACCAATCTATG 1379
Qy 752 accttagagcatcaacagctacttagaatacagaatgcaagaaacagcctctcaag 811
|||||
Db 1380 GCTTTTCGTGAGCAGCAACATCGCATATGCTCAAGGTGAAGCAGCATGGAATAAAGG 1439
Qy 812 aatctcagaagcctcctcagaagatgctgcaaggtctgtagtctcagaagctaccagaagaa 871
|||||
Db 1440 TGGGTGTGAAAAACCTGCTTTTGCCTCAGATGTGAGCGTGTCAATGCTGAAGAAGACA 1499
Qy 872 aataaagt 880
|||||
Db 1500 AATAACCGT 1508

RESULT 12
BFU75371 2315 bp DNA BCT 07-MAY-1997
LOCUS Bacteroides fragilis transposon Tn4555 transposase (tnpA) gene,
DEFINITION complete cds, and transposase (tnpB) gene, partial cds.
ACCESSION U75371
NID 92072415
KEYWORDS Bacteroides fragilis.
SOURCE Bacteroides fragilis.
ORGANISM Bacteroides fragilis.
REFERENCE 1 (bases 1 to 2315)
AUTHORS Tribble,G.D., Parker,A.C. and Smith,C.J.
TITLE The Bacteroides mobilizable transposon Tn4555 integrates by a
site-specific recombination mechanism similar to that of the
gram-positive bacterial element Tn916
JOURNAL J. Bacteriol. 179 (8), 2731-2739 (1997)
MEDLINE 97252504
REFERENCE 2 (bases 1 to 2315)
AUTHORS Tribble,G.D., Parker,A.C. and Smith,C.J.
TITLE Direct Submision
JOURNAL Submitted (18-OCT-1996) Microbiology and Immunology, East Carolina
University, Moye Blvd., Greenville, NC 27858-4354, USA
FEATURES
source 1..2315
location/Qualifiers
1..2315
/organism="Bacteroides fragilis"
/transposon="Tn4555"
/db_xref="taxon:817"
194..224
/function="transposon termini"
569..1504
/gene="tnpA"
569..1504
/gene="tnpA"
/codon_start=1
/transl_table=11
/product="transposase"
/db_xref="PID:92072417"
/translacion="KATKRCSTFCGKSEYTRSGCMORYCEACQAEKRRARVOKNNLF
KVAQPLMEIOHDEYLFESKAALMGCSROYIYKLAIGLKASRISNRAEIRRDIE
OMLEGNPHYRILPGNTSTPRKSSSLPAKREKRESEEVDFYSGEVEMLFYKO
SWLYTSAKRNHPIPCRIAGKNYSKSHIDIEFGVANDISEITDMLTEVEVDFGMP
TALRAYTRHKIPIKREYGRTRYKSHNELRRTLVNDERITYTEQVQITGLSSAN
ICHYKVKHIEIKYGVKNLLRSDEYVMAERK"
1646..2315
/gene="tnpB"
1646..2315
/gene="tnpB"
/note="integrase"
/codon_start=1
/transl_table=11
/product="transposase"
/db_xref="PID:92072416"
/translacion="MSKCTVTLRKRRIKNGTQSYSLDYPYGRDNTVMRTVEAL
GIYIFAKPANOOERDFNARMKKAVALRNQREAIENENGFPEDTKMKGDPLAYFKG
LADRNKIKMQUHYKHFQRFVNGKCTFEEDVDLCRKFMEYLLDAPQSIHTNOKLHNS
```

```
BASE COUNT 739 a 533 c 498 g 545 t
ORIGIN
Query Match 4.0%; Score 35.4; DB 36; Length 2315;
Best Local Similarity 49.2%; Pred. No. 7.5;
Matches 93; Conservative 0; Mismatches 96; Indels 0; Gaps 0;
Qy 692 gaagtagctgttagcatgtcacccaagagttctcaacagacacctgagccttaag 751
|||||
Db 1320 GAACCTGACCTTTGGAACGATGAACGCTACTATACCGTTGACGAGGTGCACCAATCTATG 1379
Qy 752 accttagagcatcaacagctacttagaatacagaatgcaagaaacagcctctcaag 811
|||||
Db 1380 GCTTTTCGTGAGCAGCAACATCGCATATGCTCAAGGTGAAGCAGCATGGAATAAAGG 1439
Qy 812 aatctcagaagcctcctcagaagatgctgcaaggtctgtagtctcagaagctaccagaagaa 871
|||||
Db 1440 TGGGTGTGAAAAACCTGCTTTTGCCTCAGATGTGAGCGTGTCAATGCTGAAGAAGACA 1499
Qy 872 aataaagt 880
|||||
Db 1500 AATAACCGT 1508

RESULT 13
MMRARG/c 1912 bp RNA ROD 09-APR-1991
LOCUS MMRARG Mouse mRNA for retinoic acid receptor gamma.
DEFINITION X15848
ACCESSION X15848
NID 953908
KEYWORDS developmental regulation; retinoic acid receptor gamma.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Euthalia; Rodentia; Sciurognath; Myomorpha; Muridae;
Murinae; Mus.
REFERENCE 1 (bases 1 to 1912)
AUTHORS Zelnit,A., Krust,A., Petkovich,M., Kastner,P. and Chambon,P.
TITLE Cloning of murine alpha and beta retinoic acid receptors and a
novel receptor gamma predominantly expressed in skin
JOURNAL Nature 339 (6227), 714-717 (1989)
MEDLINE 89295563
FEATURES
source 1..1912
location/Qualifiers
1..1912
/organism="Mus musculus"
/db_xref="taxon:10090"
/dev_stage="embryonic"
/clone_1lb="CDNA in lambda gt10"
1..1855
/note="Retinoic acid receptor gamma"
/evidence="experimental"
316..1693
/codon_start=1
/product="Retinoic acid receptor gamma"
/db_xref="PID:953909"
/db_xref="SWISS-PROT:P18911"
/translacion="MATKKEKLFAPGALGPGSGYGGAGFPFAPGALGSPFPEMLSP
SFRSGQPDLPREMAISLSEVGTSSSEKVPSPPPPRVRYKCFVCNDSSGYHY
GVSCEGCGFFRRSIOKNMYTCHRDNCLINKYTRNCOYCRQKCFEYOMSKYAV
RNDNRKKRKEVEEGSPDSELPQLEELITVSKAHOETFPCLQKGYTNSADH
RYOULGLMDKESLATAKCIKIVFAKRLPGFTGLSTADOTILKACLODILMRIC
TRYPEODTMFSDGLTLNRQMHNAGCGPLDIYAFAGOLLPLEMDTJEGLSAI
CLICDRDLDEPERKVDKIQEPLLEALRLYARRRRSPQYHFFRLMLKTTDTRGTSI
GAERAITLKMETPGMPPLIRMLENPFEDSSKPGPHRPAASSEADAPGGQGRQ
SPQPDGQ"
BASE COUNT 427 a 597 c 534 g 354 t
ORIGIN
Query Match 3.9%; Score 35; DB 13; Length 1912;
```


| | |
|-----------------|---|
| repeat_region | /db_xref="SPRMBL:015210" /translation="MKSLLLAVALGLATANSAPNATVECFVEDASGKGLAKRPGA LILRGCGPGRPDLDPELYLSVHPDPAALQAAFRKPRGAPARHCESRFPPLPAS AMAGGLTPAQCPRALDGLAMLVMSISSPVLSSILRPOPEQOEVYLITAAVYLL VTHHPAPVRLGODALLDLSEFAMPTSEASSLAPGPPGLMEORHOLGKHLIL AATPGNGMPAPAGAAVAMDDDEPMTGQTEMLTPVQFPOEGEYATLTHLP YILOGVTLLEAYVPRKYSIMPAFLARAPCARPDLICVSHFPPSGEYEMTLRG GPGGSGAKREGORMLSALRHSDGSYSGLHQPDPVTEQGAACITHPSPAS GSAEVTLEVALSGPLEDSVGLFSLALLGLFKALGMVAVYLTCKDSKKAE" /note="match: multiple ESTs; match: H95651 H95725 H95087 T02997" |
| repeat_region | 9188..9216 /note="MER5B repeat: matches 79..108 of consensus" 9432..9730 /note="AluSp repeat: matches 303..4 of consensus" 9739..10046 /note="AluSp repeat: matches 303..1 of consensus" 10228..10542 /note="AluDb repeat: matches 300..5 of consensus" 10555..10856 /note="AluSg repeat: matches 303..1 of consensus" 10924..11055 /note="AluDo repeat: matches 2..133 of consensus; incomplete repeat" |
| repeat_region | 11058..11356 /note="AluSp repeat: matches 2..303 of consensus" 11363..11474 /note="FLAM_A repeat: matches 4..133 of consensus" 11526..11657 /note="FLAM_C repeat: matches 132..1 of consensus" 11683..11966 /note="AluSp repeat: matches 294..1 of consensus" 12231..12452 /note="MER30 repeat: matches 1..230 of consensus" 12730..13036 /note="AluSg repeat: matches 293..3 of consensus" 13174..13373 /note="MIR repeat: matches 262..64 of consensus" 13378..13678 /note="AluSg repeat: matches 1..303 of consensus" 13770..14073 /note="AluSg repeat: matches 303..1 of consensus" 14084..14384 /note="AluSp repeat: matches 302..1 of consensus" 14618..14735 /note="MIR2 repeat: matches 17..146 of consensus" 14994..15182 /note="MIR repeat: matches 199..2 of consensus" |
| prim_transcript | /note="MIR repeat: matches 199..2 of consensus" 16051..16385 /note="match: 5' EST AA324236" 16785..>21386 |
| prim_transcript | /note="match: multiple ESTs; match: AA281362 AA299378 T95154 H40477 H40478; match: T95038 WA8760 WA8761 U51697 AA36474; match: R65889 H27117 N25151 D19918 R22460; match: T61710 AA404406 N35653 AA41318 R38286; match: AA303745 AA235390 AA143464 N98901 H39975; match: RA8048 AA364715 T28708 AA290846 W74553; match: AA379282 T87753 H53201 AA357767 W74659; match: H45803 H98631 AA303092 H59364 AA128601; match: R48019 T61625 AA062872 T87603 T69380; match: AA339923 AA295964 AA411317 N69910 R32928; match: AA379524 AA102006 R70438 AA335972 AA292457" 17352..17477 |
| repeat_region | /note="MER33 repeat: matches 324..197 of consensus" 17485..17603 /note="MER46 repeat: matches 12..125 of consensus" 17647..17947 /note="AluX repeat: matches 302..1 of consensus" 18000..18091 /note="MER46 repeat: matches 145..234 of consensus" 18139..18437 /note="AluSg repeat: matches 1..303 of consensus" 18454..18757 |
| repeat_region | /note="AluX repeat: matches 1..302 of consensus" 18765..18935 /note="AluDo repeat: matches 131..301 of consensus; incomplete repeat" 18947..19091 /note="MER33 repeat: matches 145..1 of consensus" 19487..19788 /note="AluX repeat: matches 302..1 of consensus" 20097..20395 /note="AluY repeat: matches 299..2 of consensus" 20119..22554 /note="putative CpG island" 20405..20699 /note="AluY repeat: matches 296..2 of consensus" 21134..21270 /note="MIR2 repeat: matches 1..139 of consensus" 24755..25052,25321..25572,25858..25761,25869..25953, 26035..26103,26318..26392,27014..27046,27164..27285, 27381..27476,27556..27667,27774..28064,28536..28650, 28767..28978) /codon_start=1 /product="RGL2" /db_xref="PID:e186796" /db_xref="PID:g2648021" /db_xref="SPRMBL:015211" /translation="MLPRRLILDTSPGCVLSPRRDPREGGCGLVVGGOE EEPEERAPVSVNDEEDGAVPTLTSNOYPLDPIYMPPRSRLRACTLEAYR HLIDTRISGTYSRNSATLATHRAFTSPALGICMLADLLESHFIDELRTTEVAL SVLSTWLSHEDGESKAGQLDRESLDTGTGAAGKGVGSADLIRNRSRVDQ APDLKPLALGDPADPTDVLVFLADLAQLTLDAELNLIPSCGLGHHRD RPGSHLSPVRAVATPTGFKAGAVSVSLATSGTEGCVITRPLPPARLLEK MIRVAEECRILNRFNSVYAVVASLOSP IHLRAAMGATRDSLRVSLQITSEED NYSOSRELLNOEVRLSPLEPHSKKAPRSSGGCVGVCITPTKDIYMLDAASDEL ENGRINFPKRKEERAVSELRLONECGYVLQDPDHIQRLQGLRPLTEQSRVSC EEPEPGSSDPAPRAVLRPTLVISQTEVLGAVGVPPLVSCDRSTGDEAPTAPAL LTRLQAQHKMWSVSLDLSALBSPSLHBPAPSHLSPASSPRRSRRASGCSPL SGCAEASGCTGCGEGSGPGASDCIRITRVOMELEDGSYKSLVTGQDAPVSR VLKNNRDSAVASEYEVLQPLGERELTIPASAVFYMDGASHDFLLRGRRSSTAT PGVSGPSASGTPSEGGGSGFPRIKATGRIARALP" |
| prim_transcript | <24790..>29426 /note="match: multiple ESTs and cDNAs; match: R37588 R25789 AA345798 C04496 D31579; match: AA400504 T32323 N31477 N80842 AA401259; match: R17716 AA401972 AA339803 C02440 AA069642; match: T08668 C03271 T08669 N99972 AA369649 D30835; match: N27045 AA400614 AA402117 AA203308 AA404233; match: R86012 AA315350 AA007622 T31165 N27365; match: AA007661 W07350 AA311078 T80173 T49124; match: T32376 AA135176 AA035367 T30397 W05771; match: N92937 T87606" |
| repeat_region | 26520..26647 /note="MIR repeat: matches 110..240 of consensus" /note="match: multiple ESTs; match: C21091 AA202721 <30147..>31460 /note="match: multiple ESTs; match: C21091 AA202721 AA409103 AA277765 AA234979; match: AA020628 AA270316 AA240162 WA88772 D25580; match: W05039 AA152842 AA359134 N74677 AA289910; complement(join(30235..30364,30630..30754,30836..30906, 31160..31223)) /codon_start=1 /product="RKE2" /db_xref="PID:e1186797" /db_xref="PID:g2648022" /db_xref="PID:g2648022" /translation="MAELIQRKLOGEYVERYOOLQDKLSKNSGKROKLEAQLTENNIVK BELALDSDSNVFKLGPVLVQKEIGEARATVGRKRDYITAEIKRYESQLDLEBROSE QORETLAQOEFORQAQKAGARGKA" |
| prim_transcript | <31845..>32746 /note="match: multiple ESTs; match: H52575 T97040" 31884..32317 /note="putative CpG island" |
| misc_feature | /note="match: CpG island DNA genomic MseI fragment 263744" complement(join(31893..33197) |


```
incomplete repeat"
repeat_region 7065..7353
/note="AluX repeat: matches 292..1 of consensus"
repeat_region 7401..7549
/note="FLAM_C repeat: matches 1..133 of consensus"
repeat_region 7555..7854
/note="AluSg repeat: matches 1..300 of consensus"
repeat_region 8081..8355
/note="AluX repeat: matches 28..298 of consensus;
incomplete repeat"
repeat_region 8384..8572
/note="AluX repeat: matches 82..268 of consensus;
incomplete repeat"
repeat_region 8573..8698
/note="AluB repeat: matches 273..146 of consensus;
incomplete repeat"
repeat_region 8716..9014
/note="AluS repeat: matches 303..1 of consensus"
repeat_region 9037..9337
/note="Alu repeat: matches 301..1 of consensus"
repeat_region 9343..9635
/note="AluX repeat: matches 294..1 of consensus"
repeat_region 9646..9946
/note="AluX repeat: matches 299..1 of consensus"
repeat_region 9948..10083
/note="Alu repeat: matches 138..1 of consensus;
incomplete repeat"
repeat_region 10125..10425
/note="AluX repeat: matches 302..2 of consensus"
repeat_region 10624..10925
/note="Alu repeat: matches 301..1 of consensus"
repeat_region 10939..11242
/note="AluS repeat: matches 303..1 of consensus"
prim_transcript <11249..>12068
/note="match: multiple ESTs; match: R99642 AA507938 C15163
C14931 W49738; match: C16651 AA659683 F22547 AA101206
AA016121; match: H39854 W52331 AA355540 AA021197 W42613;
match: AA3855141 AA017437 AA082309 AA507244 W6331; match:
AA524194 W25545 AA493910 AA583064 AA641593; match:
AA019477 AA128792 AA487370 AA487253; match: AA652667
W74358 AA599280 AA554955 AA307689; match: AA679474 W77783
AA486587 AA679458 W67984; match: AA102501 C06289 AA056113
W03710 AA150605; match: N48018 N34178 N86412 AA311933
AA071084; match: AA576022 AA580046 W05454 AA149118
AA314746; match: AA376342 AA079802 AA630792 AA487027"
complement(11465..11982)
/gene="BING3"
complement(<11465..11982)
CDS
/gene="BING3"
/codon_start=1
/pseudo
/product="CICB2046.1"
/db_xref="PID:e1273281"
12166..12460
/note="AluS repeat: matches 1..296 of consensus"
repeat_region 12462..12497
/note="18 copies of 2 mer 86 & conserved"
repeat_region 12572..12802
/note="AluX repeat: matches 1..225 of consensus;
incomplete repeat"
repeat_region 12987..13285
/note="Alu repeat: matches 301..1 of consensus"
repeat_region 13288..13578
/note="AluS repeat: matches 292..3 of consensus"
repeat_region 13635..13929
/note="AluS repeat: matches 295..1 of consensus"
repeat_region 14229..14410
/note="Alu repeat: matches 568..382 of consensus"
repeat_region 14413..14713
/note="AluX repeat: matches 1..302 of consensus"
repeat_region 14728..14813
/note="Alu repeat: matches 387..308 of consensus"
repeat_region 14941..15244
/note="AluS repeat: matches 299..1 of consensus"
repeat_region 15245..15430
/note="Alu repeat: matches 117..302 of consensus;
incomplete repeat"
repeat_region 15456..15564
/note="MER42C repeat: matches 1043..1148 of consensus"
repeat_region 15583..15749
/note="FRAM repeat: matches 1..166 of consensus"
repeat_region 15783..16044
/note="AluX repeat: matches 34..302 of consensus;
incomplete repeat"
repeat_region 16068..16156
/note="MER42C repeat: matches 1193..1295 of consensus"
repeat_region 16147..16349
/note="MLTIC repeat: matches 466..265 of consensus"
repeat_region 16350..16373
/note="12 copies of 2 mer 100 & conserved"
repeat_region 16375..16542
/note="AluX repeat: matches 286..124 of consensus;
incomplete repeat"
repeat_region 16543..16835
/note="AluS repeat: matches 300..1 of consensus"
repeat_region 16839..16971
/note="AluS repeat: matches 133..1 of consensus;
incomplete repeat"
repeat_region 16991..17195
/note="MLTIC repeat: matches 246..27 of consensus"
repeat_region 17206..17448
/note="L1PA11 repeat: matches 911..655 of consensus"
repeat_region 17471..17773
/note="AluS repeat: matches 300..1 of consensus"
repeat_region 17783..18080
/note="Alu repeat: matches 301..1 of consensus"
repeat_region 18108..18148
/note="MLTIB repeat: matches 41..1 of consensus"
repeat_region 18168..18295
/note="MER42C repeat: matches 1414..1538 of consensus"
repeat_region 18700..18867
/note="Alu repeat: matches 291..124 of consensus;
incomplete repeat"
repeat_region 18868..19164
/note="Alu repeat: matches 297..1 of consensus"
repeat_region 19165..19299
/note="Alu repeat: matches 135..1 of consensus;
incomplete repeat"
repeat_region 19376..19650
/note="Alu repeat: matches 6..279 of consensus;
incomplete repeat"
repeat_region 19654..19946
/note="AluS repeat: matches 3..294 of consensus"
repeat_region 20198..20505
/note="AluX repeat: matches 302..1 of consensus"
repeat_region 20529..20560
/note="16 copies of 2 mer 88 & conserved"
repeat_region 20668..20971
/note="AluX repeat: matches 1..302 of consensus"
repeat_region 22058..22357
/note="AluS repeat: matches 303..1 of consensus"
repeat_region 22475..22777
/note="AluX repeat: matches 1..302 of consensus"
repeat_region 22863..23083
/note="AluS repeat: matches 296..73 of consensus;
incomplete repeat"
repeat_region 23084..23387
/note="AluX repeat: matches 302..1 of consensus"
repeat_region 23398..23511
/note="Alu repeat: matches 122..9 of consensus;
incomplete repeat"
repeat_region 23525..23824
/note="AluS repeat: matches 299..1 of consensus"
repeat_region 23827..24125
/note="AluX repeat: matches 297..1 of consensus"
repeat_region 24128..24422
```



```

repeat_region /note="AluX repeat: matches 296. .1 of consensus"
24490. .24801
/note="AluSg repeat: matches 1. .295 of consensus"
repeat_region 25092. .25405
/note="AluY repeat: matches 297. .1 of consensus"
repeat_region 25459. .25581
/note="AluO repeat: matches 3. .125 of consensus;
incomplete repeat"
repeat_region 25584. .25880
/note="AluSx repeat: matches 299. .1 of consensus"
repeat_region 25881. .26042
/note="AluO repeat: matches 115. .277 of consensus;
incomplete repeat"
repeat_region 26062. .26361
/note="AluSc repeat: matches 299. .1 of consensus"
repeat_region 26374. .26423
/note="25 copies of 2 mer 88 & conserved"
prim_transcript <28015. .>32406
/note="match: multiple ESTs; match: AA129657 AA386922
245843 AA405989 T79567; match: W39776 H13819 H49831 W02333
H77913 H73773; match: H34947 AA613090 AA075116 AA383223
AA493034; match: AA317661 C03211 AA085057 AA301687 W45308;
match: AA356615 W51682 AA405738 AA355087 AA129682; match:
AA355781 C00363 AA374951 AA535586 Z41479; match: H73021
T79477 AA669990 R02432 H74158 H45863; match: AA362200
AA308321 AA075005 R48034 T36090; match: AA312483 AA234490
AA644233 N73287 T30378; match: AA523436 AA312767 H47926
AA317476 AA641449; match: AA380467 AA365654 AA306929
AA305678"
29039. .32221
/note="Daxx"
gene /gene="Daxx"
301n(29039. .29245,29397. .30228,30373. .30584,30740. .30953,
31110. .31584,31745. .31967,32162. .32221)
/gene="Daxx"
/codon_start=1
/product="C1CB2046.2"
/db_xref="PID:e1273283"
/transcript="MATANSIIVLDDDDDEDAAOOPSPHPLPNAASPGAEAPSSSE
HAGRSSSSGGKCKYKLENEKLEFEELFEELKOTAPKPVLPVLPVLPVLPVLPVLPV
AERCNILSVLSARSRPAKIYIINELTVLKAHSAKKLNLAIPATTSNEDSGNP
PTHSLDPTNAENTASQSPRTGSRKQIRLEQLALVAEIRLQEKELDSELDPP
DSAYLOEARLKRLLRFGRLCELDKDSITGVIVORIPYRGRPEVNRRIERLIN
KGGPDPDYGVILRAVEKAAAHSLGSLRQOOLAAQOAFVGVIRLQERHLDIY
KNGCHLTDYRPGVDALDPVYARLRREMSLRMSRLDEIVSKYMLDOKSEEGEK
KRRARLOGSSHSADTPESALDSGEGPSMAISGCGSARAEVDDDDDESDDEEER
EEEEERATDSEEREDLEQMOBQDEDEDEDEDEEAAGRGDKSPWSLQISNEKNL
EFGKQISRSSEGOQNGRIVSPSLISEEPLAPSSIDAESNGEQPELTLSEESPVSQ
FLELEALPLDTPSSVETDISSSRKQSEEPFTVLNGAGMVSSTFNGVSPHMGD
SPPCKSKRERKQSGPLGNSYVERORSYHEKNGKCTCLPSPPSLASLAPVADS
STRVDSPHGIVTSSLCIPSPARLISOTPHSQPPRPTCTSVATQCDPELIIYLSDD
"
repeat_region 32523. .32739
/note="Aluub repeat: matches 302. .88 of consensus;
incomplete repeat"
repeat_region 33739. .33907
/note="MIR repeat: matches 20. .206 of consensus"
gene 34049. .35953
/gene="BING1"
34049. .35953
/gene="BING1"
/gene="BING1"
/codon_start=1
/product="C1CB2046.3"
/db_xref="PID:e1273284"
/db_xref="PID:g2961408"
/transcript="MEPSPLSPGALPLPLSLAPPLPLPAAVVHVSPFEVTSALL
ESLNOQRLGOLCGLSIRVQGEFFRAHRAVLAASPYFDQVLLKMTSLSLSPMDP
GAFFYLASAYTGLSMAADIYNFLVGSVLOMHIIDKCTELLREGASATTTT
AAATSVTPGAGVPGSGGTAVPATGASARSHASRSASNSQSSSNYSFPRSTDFS
SSQBEPAASAVSGSERRGGGPTFPAPVYGSGGATIGKLLLEADELDCDDGGDGAAY
PAGGLRPTTPPSLIPQKHWTYVRKGNCPAPTPVLPQDDPLEEEEEEEDLVLTED
DDEDELGSSRVVGGGPEATLSISVRTLSEPPDGGEBGVNFCSSNDGFEYEGGAE
VAGLDSGGPTSPSSYAPSHPPRLPLDQGNQIIVFPSSSSSSSSSQAPGQPGNGAE

```

```

prim_transcript <34893. .>35752
/gene="BING1"
/note="match: multiple ESTs; match: AA085230 AA056630
AA309450 N43919 AA468696; match: R53654 R36276 T07529
H15337 W78260 AA4051601"
prim_transcript <35824. .>36547
/note="match: multiple ESTs; match: T78477 T78559 AA359884
D82810 H34158 H15719; match: AA465000 AA062251 T07530
T03652 R53543 N34845; match: AA056605 N23868 T27481
AA455280 AA018565; match: T18609 R49563 W74632 W94380
N29958"
prim_transcript <36919. .37233
/note="match: 5' EST AA280744 clone IMAGE:711440"
36924. .37748
/gene="tapasin"
301n(36924. .36960,37101. .37271,37488. .>37748)
CDS /gene="tapasin"
/note="continues in 297184"
/codon_start=1
/product="C1CB2046.4"
/db_xref="PID:e1273285"
/db_xref="PID:g2961409"
/transcript="MKSILIAVALGATATASAPAVTECFVENDASGGLAKRPGA
LLIRGCGPPRPRLDPELYISVDPRGALQAAARRRPRGAPARHCESRVLPAS
AKWASGLTPAQCPRALDQAWLWVSISSVLSLSLRQPPQEPVLIYAT"
prim_transcript <37882. .>38774
/note="match: multiple ESTs; match: H95651 H95725 H95087
T02997"
repeat_region 39072. .39100
/note="MER5B repeat: matches 79. .108 of consensus"
repeat_region 39316. .39614
/note="AluSp repeat: matches 303. .4 of consensus"
repeat_region 39623. .39872
/note="AluSg repeat: matches 303. .59 of consensus;
incomplete repeat"
BASE COUNT 9648 a 9448 c 10468 g 10308 t
ORIGIN
Query Match 3.9%; Score 34.8; DB 11; Length 39872;
Best Local Similarity 50.6%; Pred. No. 16;
Matches 84; Conservative 0; Mismatches 82; Indels 0; Gaps 0;
QY 44 tggtagagccatgagcactcaagtcctacctaacacagagcgtgcttcgtgccca 103
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36788 TGGGGAACACCGGTCGCCAGGACACCTTACCTAACACAGCGGGGATTCCACGCC 36847
QY 104 ttctctctactcaactcaactgcatgtaagtgagtgcggaacggagattcgagcagc 163
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36848 CTCATGCCGCCCTTTGGAGGAAGTGAAGTGAAGAGAGGAAGAGAGCTTCATGCTG 36907
QY 164 aggaattcaagatgcatgctggaactggtcctctgcaaacggtg 209
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 36908 AGGAGTCCGACGCGCATGAAGTCCGTCTGCTCTCTGCTGTG 36953

```

Search completed: May 14, 1999, 03:31:54
 Job time: 5533 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 11:42:04 ; Search time 129.53 Seconds
(without alignments)
1297.006 Million cell updates/sec

Title: US-09-212-270-8

Perfect score: 893
Sequence: 1 tccggcgccggcgggcgagga.....taaaagtgtctatacccta 893

Scoring table: IDENTITY_NUC

Searched: 240622 seqs, 94065609 residues

Database : N_Geneseq_34:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|------------|---------------------|
| 1 | 447.2 | 50.1 | 1704 | 1 V33361 | Nucleotide sequenc |
| 2 | 447.2 | 50.1 | 1706 | 1 V33362 | Nucleotide sequenc |
| 3 | 284.8 | 31.9 | 371 | 1 V11422 | Human secreted pro |
| 4 | 69.8 | 7.8 | 181 | 1 V11423 | Human secreted pro |
| 5 | 35 | 3.9 | 1912 | 1 Q10388 | Encodes skin-speci |
| 6 | 35 | 3.9 | 1913 | 1 T05202 | Murine retinoic ac |
| 7 | 34 | 3.8 | 110000 | 1 V30458_2 | Continuation (3 of |
| 8 | 33.8 | 3.8 | 3013 | 1 Q94328 | Rabbit telencephal |
| 9 | 33 | 3.7 | 2810 | 1 N70218 | LEU2 gene of Y. lip |
| 10 | 33 | 3.7 | 2810 | 1 Q83184 | Yarrowia lipolytic |
| 11 | 31 | 3.5 | 5822 | 1 V62933 | Human mdia Rho tar |
| 12 | 31 | 3.5 | 5822 | 1 V62934 | Human mdia Rho tar |
| 13 | 30.6 | 3.4 | 3471 | 1 Q27167 | Delta-endotoxin ge |
| 14 | 30.6 | 3.4 | 3471 | 1 Q28940 | B. thuringiensis ps |
| 15 | 30.6 | 3.4 | 3471 | 1 Q30821 | Toxin 50C. New Bac |
| 16 | 30.6 | 3.4 | 3471 | 1 Q38653 | Bt isolate Ps50C c |
| 17 | 30.6 | 3.4 | 3471 | 1 Q51704 | Bacillus thuringie |
| 18 | 30.6 | 3.4 | 3471 | 1 T43222 | Antiscarab pest to |
| 19 | 30.6 | 3.4 | 674 | 1 V15992 | Streptococcus pneu |
| 20 | 30.4 | 3.4 | 7760 | 1 V52196 | Polymer clone pY12 |
| 21 | 30 | 3.4 | 1878 | 1 V41379 | Scyrtococcus pneu |
| 22 | 29.8 | 3.3 | 13875 | 1 Q20240 | Murine NF-KB recep |
| 23 | 29.8 | 3.3 | 13875 | 1 Q20240 | Murine NF-KB recep |
| 24 | 29.6 | 3.3 | 3366 | 1 Q84803 | Human centromeric |
| 25 | 29.4 | 3.3 | 3366 | 1 Q84803 | Spinocherebellar at |
| 26 | 29.4 | 3.3 | 3366 | 1 Q84803 | Encodes Porcine pa |
| 27 | 29.4 | 3.3 | 3366 | 1 Q84803 | Myobacterium tube |
| 28 | 29.2 | 3.3 | 49272 | 1 V35010 | Sarcosine oxidase. |
| 29 | 29.2 | 3.3 | 565 | 1 Q24075 | HCV envelope regio |
| 30 | 29 | 3.2 | 1394 | 1 Q45175 | Sequence encoding |
| 31 | 29 | 3.2 | 2791 | 1 Q99776 | Phosphomycin diosy |
| 32 | 29 | 3.2 | 1977 | 1 T16507 | Rice HYA cDNA part |
| 33 | 28.8 | 3.2 | 2098 | 1 Q66177 | Seven transmembran |
| 34 | 28.8 | 3.2 | 3754 | 1 T28528 | S. pneumoniae dete |
| 35 | 28.8 | 3.2 | 137507 | 1 V19941 | KSHV long unique c |
| 36 | 28.8 | 3.2 | 2098 | 1 V19355 | Human R2 seven tra |
| 37 | 28.6 | 3.2 | 940 | 1 Q06394 | Exon 2 of porcine |
| 38 | 28.6 | 3.2 | 755 | 1 Q32449 | HCV core-envelope |
| 39 | 28.6 | 3.2 | 1258 | 1 Q32450 | HCV core-envelope |
| 40 | 28.6 | 3.2 | 1554 | 1 Q32451 | HCV core-envelope |
| 41 | 28.6 | 3.2 | 1258 | 1 Q32452 | HCV core-envelope |
| 42 | 28.6 | 3.2 | 1037 | 1 Q58450 | Hepatitis C virus |
| 43 | 28.6 | 3.2 | 2033 | 1 Q64913 | Hepatitis C virus |

ALIGNMENTS

| RESULT | ID | Score | Query Match | Length | DB ID | Description |
|--------|--------|-------|-------------|--------|-------|--|
| 1 | V33361 | 44 | 28.6 | 3.2 | 3366 | 1 Q84803 |
| | V33361 | 45 | 28.6 | 3.2 | 2033 | 1 Q86788 |
| | V33361 | | | | | Spinocherebellar at |
| | V33361 | | | | | Hepatitis C virus |
| | V33361 | | | | | Standard; cDNA to mRNA; 1704 BP. |
| | V33361 | | | | | 02-DEC-1998 (first entry) |
| | V33361 | | | | | Nucleotide sequence of human alpha-OAF065. |
| | V33361 | | | | | Human; alpha-OAF065; stroma cell; antibody; inflammatory; |
| | V33361 | | | | | cytokine-mediated disease; rheumatism; ulcerative colitis; ss. |
| | V33361 | | | | | Homo sapiens. |
| | V33361 | | | | | Location/Qualifiers |
| | V33361 | | | | | CD5 |
| | V33361 | | | | | 45..1298 |
| | V33361 | | | | | /tag= a |
| | V33361 | | | | | /product= "human alpha-OAF065 protein" |
| | V33361 | | | | | 45..119 |
| | V33361 | | | | | /tag= b |
| | V33361 | | | | | 120..1295 |
| | V33361 | | | | | /tag= c |
| | V33361 | | | | | /transl_except= (pos:711..713, aa= Pro) |
| | V33361 | | | | | /transl_except= (pos:714..716, aa= Arg) |
| | V33361 | | | | | W09838304-A1. |
| | V33361 | | | | | 03-SEP-1998. |
| | V33361 | | | | | 26-FEB-1998: J00799. |
| | V33361 | | | | | 27-FEB-1997: JP-043143. |
| | V33361 | | | | | (ONOT) ONO PHARM CO LTD. |
| | V33361 | | | | | Fukushima D, Konishi M, Tada H; |
| | V33361 | | | | | WPI: 98-481205/41. |
| | V33361 | | | | | P-PSDB: W70386. |
| | V33361 | | | | | Membrane polypeptide expressed by human stroma cells, and antibodies |
| | V33361 | | | | | recognizing it--for treatment of inflammatory and other |
| | V33361 | | | | | cytokine-mediated diseases. |
| | V33361 | | | | | PS Claim 5; Pages 31-32; 54p; Japanese. |
| | V33361 | | | | | CC This is the nucleotide sequence of the human alpha-OAF065, used in |
| | V33361 | | | | | CC the method of the invention. The process involves the use of peptides |
| | V33361 | | | | | CC expressed by stroma cells, and its antibodies are used for in the |
| | V33361 | | | | | CC prevention and treatment of inflammatory and other cytokine-mediated |
| | V33361 | | | | | CC diseases such as rheumatism, ulcerative colitis. |
| | V33361 | | | | | Sequence 1704 BP; 429 A; 426 C; 430 G; 419 T; |
| | V33361 | | | | | Query Match |
| | V33361 | | | | | 50.1%; Score 447.2; DB 1; Length 1704; |
| | V33361 | | | | | Best Local Similarity 81.6%; Pred. No. 2e-129; |
| | V33361 | | | | | Matches 529; Conservative 0; Mismatches 118; Indels 1; Gaps 1; |
| | V33361 | | | | | 32 aataacacgcttggtgagagccatgcaactgaagctcctacacagagcgtg 91 |
| | V33361 | | | | | 22 AATAAATACATTGTGTAAGAAAGATGCTTTAAATGCTACAGAAAGAAATACGT 81 |
| | V33361 | | | | | 92 tcttcgctgcaactccttcctacacacgctgcaatgtaagtgaattgcgaacggag 151 |
| | V33361 | | | | | 82 TTTTCTACTCTTATTAGTATTCTAGGCTATTGTGATGTAAGTACTTGGAAGAGAG 141 |
| | V33361 | | | | | 152 attgacggcagcaggaattcaagatcgatcgtgaactgtgctcctcgaacagtgcg 211 |
| | V33361 | | | | | 142 ACTGTGACACGACGAAGATTCAGGATCGCTGGAACCTGTCCCTGCACAGTGTG 201 |
| | V33361 | | | | | 212 gacctgcatgaggtgtctcaagaagtgtgctcgtgctatgggaggaagcagtg 271 |
| | V33361 | | | | | 202 GGCACAGCATGAGGTGTCTCAAGATGTGCTTCGGCTATGGGAGAGATGACAGTGTG 261 |
| | V33361 | | | | | 272 tgcctcgtcagcgcgcagcaggttcaagaagactgggtttccagaagtgtgaacatgtg 331 |
| | V33361 | | | | | 262 TGACGTGCGCGCTGCACAGAGTCAAGAGACTGGGGCTCCAGAAATGCAACCTGTGTC 321 |
| | V33361 | | | | | 332 cggactgtgctgctggtgagacgctttcagaagggcgaactgtcaccacacagcgtgagcgtg 391 |
| | V33361 | | | | | 322 TGGACTGCGCAGTGTGTAAGACCGCTTTCAGAAAGCAATATTGTCAGCCACAGTATGCCA 381 |

| | | | |
|----|-----|---|-----|
| OY | 392 | tctgtgaggagactgcctctgcagagattttacccgaaagaaacaaacttggtgtttcaagaca | 451 |
| Db | 382 | ttctggggggactgcgttggccaggattttttatggaaacagaaactttctggcttttaagaca | 441 |
| OY | 452 | tggagttgtgtgccctgcgcgagaaaccacactccctccactacgaacacactgtacagacaag | 511 |
| Db | 442 | tggagttgtgtgcttctgtggagacccctccctccttttaggaacccgacgtgtgcacacaag | 501 |
| OY | 512 | tgaacctgttgaagttcttcctccaccgctctccagccctccggagacacggtgtgtctcgg | 571 |
| Db | 502 | tcAACCTGtGAAGATCGGtGTCACAGGcCTCCAGGCCACAGGGGACACAGGGcCTGGGTGGCG | 561 |
| OY | 572 | tcaatctgcagttgctcttggccacagcgtgtgtctcg-ctgtgtcaatctgtgtgtcatctact | 630 |
| Db | 562 | TTATCTGCAGcGCTCGGGCACCGtGTCGTCGTCGCTGCTGTCATCTCTGTGTCATCTATT | 621 |
| OY | 631 | gcaaaagcagttctcatgtgagaagaacaccacagctcttaagctcccatccc | 678 |
| Db | 622 | GTAACAGACAGCTTTATGTGAGAGAAACCCAGCTGGTCTCTGGGGTCAC | 669 |

```

RESULT      2
ID           V33362
V33362      V33362 standard; cDNA to mRNA; 1496 BP.
AC           V33362;
DE           02-DEC-1998 (first entry)
DE           Nucleotide sequence of human beta-OAF065.
KW           Human; beta-OAF065; stroma cell; antibody; inflammatory;
KW           cytokine-mediated disease; rheumatism; ulcerative colitis; ss.
OS           Homo sapiens.
FH           Key
FT           CDS
FT           Location/Qualifiers
FT           45..1316
FT           /*tag= a
FT           /*product= "human beta-OAF065 protein"
FT           45..119
FT           /*tag= b
FT           120..1313
FT           /*tag= c
FT           /*transl_except= (pos:711..713, aa= Pro)
FT           /*transl_except= (pos:714..716, aa= Arg)

WO9838304-A1.
PD           03-SEP-1998.
PF           26-FEB-1998: JP-043143.
PR           27-FEB-1997: JP-043143.
PA           (ONOD ) ONO PHARM CO LTD.
PI           Fukushima D, Konishi M, Tada H;
PI           WPI; 98-481205/41.
DR           P-PSDB; W70387.
DR           Membrane polypeptide expressed by human stroma cells, and antibodies
PR           recognising it--for treatment of inflammatory and other
PR           cytokine-mediated diseases.
PS           Disclosure; Pages 40-41; 54pp; Japanese.
CC           This is the nucleotide sequence of the human beta-OAF065, used in
CC           the method of the invention. The process involves the use of peptides
CC           expressed by stroma cells, and its antibodies are used for in the
CC           prevention and treatment of inflammatory and other cytokine-mediated
CC           diseases such as rheumatism, ulcerative colitis.
SQ           Sequence 1496 BP; 388 A; 360 C; 372 G; 376 T;

```

| | | | | |
|-----------------------|-----------------|---------------------|-----------|--------------|
| Query Match | 50.1%; | Score 447.2; | DB 1; | Length 1496; |
| Best Local Similarity | 81.6%; | Pred. No. 1.9e-129; | | |
| Matches 529; | Conservative 0; | Mismatches 118; | Indels 1; | Gaps 1; |

[illegible]

| | | | |
|----|-----|--|-----|
| QY | 152 | atggaagcacaacagaaatctcaagatcgatctgnaactctgtccctctgaaacagtcg | 211 |
| Db | 142 | ACTGTAGACAGACGAATAATCCAGGATGGCTGGAAACTGTGTCTCTGCACACAGTGT | 201 |
| QY | 212 | gacctgcatactgagttgtctccaagaatgtggtcttcggtctatgaggaagatgcacgttg | 271 |
| Db | 202 | GCCCGGCGATGAGATTGCTTAAGAAATGTGGCTTCGGCTATGGGGAGATGCACAGTGC | 261 |
| QY | 272 | tgcccttcgaagccgacccggttcaaggaagacttgggtttcccgaaatgtgaagccatgtg | 331 |
| Db | 262 | TGACCTGCCGGGTGCACAGGTTCAGGAGAGATGGCTTCACAAATATGCAAGCCCTGTC | 321 |
| QY | 332 | cggactctgcctgctgtgagaccgcttccagaaagccaaactctcacacaccagtcgatctg | 391 |
| Db | 322 | TGGACTGCGCAGTGGTGTAACCGCTTTCAGAAAGCAAAATGTTTCAGCACACAGTATGCCA | 381 |
| QY | 392 | tctgtggggactgcctgcgcgaagatttcaacggaagaccaaactgtgtgtttccaagaca | 451 |
| Db | 382 | TCTGGGGGAGCTGCTGCACAGATTTTATAGAAAGACGAACCTTGTGGCTTTCAMACA | 441 |
| QY | 452 | tggagctgtctccctcgcggaagaccacactctccctcagaaaccacactgtctacagcaag | 511 |
| Db | 442 | TGGATGTGTGCTTTGTGTGGAAACCCCTCTCTCTCTTACGAACCCCACTGTGTCCAGCAAG | 501 |
| QY | 512 | tgaacctgttgaagatctctctccacacctctccagccttcgggacagggcgtgtgctgcg | 571 |
| Db | 502 | TCCAACTCGTAAAGATCCGCTCCACAGGCTCCACACCGGAGACACGGCGCTGGCTCCG | 561 |
| QY | 572 | tcaatctgcacatcctctggcacaagctgtcgtctcg-cctgcatcctgttgttctact | 630 |
| Db | 562 | TTATGTGCAGCGCTGTGGCCACCGTCCGCTGGCCCTGTCTATCTCTGTGTCTATATT | 621 |
| QY | 631 | gcaagaagcagttcatctggagaagaaccacagctgtlaagctccatccc | 678 |
| Db | 622 | GTAAAGACACTTTATGTGAGAGAAACCCACTGCTGTCTGGCTGCAC | 669 |

RESULT 3
V11422
ID V11422 standard; cDNA; 371 BP.
AC V11422;
DR 23-JUL-1998 (first entry)
DE Human secreted protein clone AX92_3 cDNA 5'-end.
KM Secreted protein; prevention; treatment; gene therapy; ds.
OS Homo sapiens.
PN WO9801554-A2.
PD 15-JAN-1998.
PE 07-JUL-1997; U11876.
PR 09-JUL-1996; US-677231.
PA (GENY) GENETICS INST INC.
PI Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM,
PI Merberg D, Racie LA, Spaulding V, Treacy M;
DR WPI: 98-110230/10.
DR P-PSDB: M58844.
PT Secreted proteins and polynucleotides encoding them - useful to
PS prevent, treat and ameliorate medical conditions
PT Claim 15; Page 57; 93pp: English.
CC V11422-V11424 encode fragments of a novel secreted protein derived from
CC clone AX92_3 which was isolated from a human adult testes cDNA library.
CC The protein can be used to prevent, treat or ameliorate a medical
CC condition, while the polynucleotides can be used for gene therapy.
SQ Sequence 371 BP; 83 A; 96 C; 107 G; 82 T;

| | | | | |
|-----------------------|-----------------|--------------------|-----------|-------------|
| Query Match | 31.9%; | Score 284.8; | DB 1; | Length 371; |
| Best Local Similarity | 85.2%; | Pred. No. 3.1e-79; | | |
| Matches 316; | Conservative 0; | Mismatches 55; | Indels 0; | Gaps 0; |

182 cgggaactgtgtctctgcgaacagtcgcagctggcatgaattgtccaagaaatgt 241
 |||||
 1 cgggaactgtgttccctgcgaacagtcgtggccagcagatgaattgtcgaagaaatgt 60
 242 gcttcgcctatcgggaagatgcacagttgttgccttcgaagccgacacggttcaagaag 301

```

Db      61  GCTTGGCTATGGGAGATGACAGTGTGTACCTGCCGGGTGCACAGCTTCAAGAGG 120
Qy      302 actgagggttcacagaagtgttaagccatgtgcgactgtgcgtgtgaacgccttcaga 361
Db      121  ACTGGGCTTCCAAAATGCAAGCCCTGTCTGTGACTGTGCAGTGTGTAACCCCTTTCAGA 180
Qy      362  gggccaactgtctcacacacccagtgatgtctgtctgcgggagactgcgtccagagatttacc 421
Db      181  AGGCAATTTGTTACACCCAGTATGATCCATCTGGGGGACTGCTGCCANAGATTATTATA 240
Qy      422  ggaagaccacaaactgtgtgttttcaagaacatgagtggtgtccctgcggagagaccacctc 481
Db      241  GGAAGACAAACTTGTCCGCTTTCAAAACATGAGAGTGTGTGCTTGTGGAACCCCTCTC 300
Qy      482  ctccctacgaacacacactgttaccagaagtgagtgacctgttgaagatctctccaccgct 541
Db      301  CTCCTTAGCAACCCGACCTGTGCACGAAGTCAAGTCAACCTGCTGAATTCNGTCCACGGCCT 360
Qy      542  ccagccctcgg 552
Db      361  CCAGCCACGG 371

RESULT 4
V11423
ID      V11423 standard: cDNA; 181 BP.
AC      V11423;
DE      23-JUN-1998 (first entry)
KW      Human secreted protein clone AX92.3 cDNA internal fragment.
OS      Homo sapiens.
PN      W09801554-A2.
PF      07-JUL-1997; U11876.
PD      15-JAN-1998.
PR      09-JUL-1996: US-677231.
PA      (GENY ) GENETICS INST INC.
PI      Bowman M, Evans C, Jacobs K, LaVallie ER, McCoy JM,
DR      Merberg D, Racie LA, Spaulding V, Treacy M;
P-PSDB: W58844.
PT      Secreted proteins and polynucleotides encoding them - useful to
PT      prevent, treat and ameliorate medical conditions
PS      Claim 13, Page 57; 93pp: English.
CC      V11422-V11424 encode fragments of a novel secreted protein derived from
CC      clone AX92.3 which was isolated from a human fetal brain cDNA library.
CC      The protein can be used to prevent, treat or ameliorate a medical
CC      condition, while the polynucleotides can be used for gene therapy.
SQ      Sequence 181 BP; 43 A; 47 C; 34 G; 57 T;

Query Match      7.8%; Score 69.8; DB 1; Length 181;
Best Local Similarity 86.5%; Pred. No. 1.6e-12;
Matches 77; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy      413  gatttaccggaagacaaactgtgtgttttcaagaacatgagtggtgtccctgcggag 472
Db      88  GATTTTATAGAGAGACGAACCTGTGCGCTTTCAGACATGAGTGTGCTTGTGAG 147
Qy      473  acccaccctccctcagaacacactgt 501
Db      148  ACCCTCTCTCTTACGACACCGCACTCT 176

RESULT 5
O10388/C
ID      O10388 standard: cDNA; 1912 BP.
AC      O10388;
DE      11-APR-1991 (first entry)
DE      Encodes skin-specific murine Retinoic Acid Receptor-gamma.
KW      mouse; retinoic acid receptor; mRAR-gamma; embryogenesis; ss.
OS      Mus musculus.
FT      key      Location/Qualifiers

```

```

FT      cds      316..1892
FT      FT      /*tag= a
FT      PN      EP-41323-A.
FT      PD      06-FEB-1991.
FT      PF      29-JUN-1990; 112469.
FT      PR      30-JUN-1989; US-374690.
FT      PR      29-MAR-1990; US-502140.
FT      PA      (INRM ) INSERM INST NAT SANTE.
FT      PA      (SOUT ) LES LABS SOUTIB SA.
FT      PI      Chambon P, Kastner P, Krust A, Petkovich M, Zient A;
FT      PI      Leroy P, Mendelsohn C, Staub A;
DR      WPI: 91-038271/06.
DR      P-PSDB: R10549.
PT      Novel human and mouse retinoic acid receptors - encode proteins
PT      used to assay for agonists and antagonists
PS      Disclosure: Fig 1; 33pp: English.
CC      An 11.5 day-old total mouse embryo lambda gt10 cDNA library was
CC      screened with labelled human RAR-alpha and -beta probes. Positive
CC      clones were isolated, mapped and sequenced. A set of clones was
CC      identified having a lesser degree of homology with either RAR-alpha
CC      or -beta than murine RAR-alpha or mRAR-beta clones, respectively.
CC      The cDNA-deduced amino acid sequence of this third set of clones
CC      was obviously related to mRAR-alpha and mRAR-beta. This new member
CC      of the mouse RAR subfamily was designated RAR-gamma. The
CC      mRAR-gamma cDNA clone was used to clone 7 different murine
CC      RAR-gamma cDNA isoforms, generated by alternative splicing of at
CC      least 7 exons. The isoforms were found to differ in their 5'-UTR
CC      sequences and in their N-terminal A region, which is known to be
CC      important for differential trans activation by other nuclear
CC      receptors.
CC      See also R10547-8, Q10389 and Q10405-8.
SQ      Sequence 1912 BP; 427 A; 597 C; 534 G; 354 T;

```

```

Query Match      3.9%; Score 35; DB 1; Length 1912;
Best Local Similarity 47.5%; Pred. No. 0.37;
Matches 104; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

Qy      146  ccggaagatgcagcagcaggaatccaagatcgatcgatcgatgctcctcgcaaac 205
Db      223  CTTGGCTTGGAGAGCTCTGCTCCCACTGCTGCTGTCTCAAGTGGCTTTCAGAGATC 174
Qy      206  agtgcgacctgcagatgagtgatgttcacagaagtgtgcctgcgtatgaggagatgcac 265
Db      173  CCTTGGCTTGCAGAGATGCTCTCTCACTACTGTGGCCAGCGTGAAGTGAAGACTG 114
Qy      266  agtgtgtcctgcagcagcagcagtggttcaagaagactggtgttccagaagtgtgaac 325
Db      113  GTCGAATCTGCTGAGCGGAGCGGGGGAAGAGCCGCGAGCCCGAGTCCCGGC 54
Qy      326  catgtgcgactgtgcgtgtgtgaacgcttcagagg 364
Db      53  GTGGGGCGGTCTCGGGGATGAGCACCGCGGTGGGGGG 15

RESULT 6
T05202/C
ID      T05202 standard: cDNA; 1913 BP.
AC      T05202;
DE      13-JUN-1996 (first entry)
DE      Murine retinoic acid receptor gamma cDNA.
KW      Retinoic acid receptor; mouse; RAR; RAR-gamma; transcription factor;
OS      Mus musculus.
FT      key      Location/Qualifiers
FT      FT      cds      316..1692
FT      FT      /*tag= a
FT      FT      /product= retinoic acid receptor gamma protein
FT      FT      misc_feature 316..498
FT      FT      /*tag= b
FT      FT      /note= "hRAR-gamma A region"
FT      FT      misc_feature 499..582

```

```

FT FT /tag= C
FT FT /note= "hRAR-gamma B region"
FT FT misc_feature 583..780
FT FT /tag= d
FT FT /note= "hRAR-gamma C region"
FT FT misc_feature 781..918
FT FT /tag= e
FT FT /note= "hRAR-gamma D region"
FT FT misc_feature 919..1578
FT FT /tag= f
FT FT /note= "hRAR-gamma E region"
FT FT misc_feature 1579..1689
FT FT /tag= g
FT FT /note= "hRAR-gamma F region"

EP-683227-A1.
PD 22-NOV-1995.
PF 29-JUN-1990; 110924.
PR 30-JUN-1989; US-374690.
PR 29-MAR-1990; US-502140.
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
PA (BRIM ) INST RECH SOUTIB.
PA (BRIM ) LES LAB SOUTIB SA.
PI Chandon P, Kastner P, Krust A, Leroy P, Mendelsohn C;
PI Petkovich M, Stauba, Zelent A;
PI WPI: 95-394349/51.
PI P-PSDB: R84726.
PT New cDNA encoding the human retinoic acid receptor gamma - useful
PT for identifying agonists and antagonists of RAR-gamma and detecting
PT expression in normal and diseased tissue
PS Example 1; Fig 1; 34pp; English.
CC This sequence represents the cDNA encoding the mouse retinoic acid
CC receptor gamma (RAR-gamma). This sequence was used in the isolation of
CC the human RAR-gamma sequence. RAR-gamma is a retinoic acid inducible
CC transcription factor, and is expressed specifically in the skin where it
CC is the predominant RAR. The C and E regions of the protein are the most
CC highly conserved regions through the RAR's of a species. The C region is
CC the DNA binding domain, and the E region is the ligand binding domain.
CC The B region is also well conserved within species. The N and C terminal
CC regions of segment D are conserved, whereas the central segment of this
CC region is not. The D region is thought to act as a hinge region.
CC Regions A and F are not conserved within a given species. The A, B, D
CC and F regions are conserved across species, suggesting that they have
CC specific functions, but are not required for the RAR's specific
CC physiological roles. The protein is used to detect competitive and
CC non-competitive agonists and antagonists. This sequence can be used to
CC detect receptor expression in normal and diseased tissues and cultures.
SQ Sequence 1913 BP; 427 A; 600 C; 531 G; 355 T;

Query Match 3.9%; Score 35; DB 1; Length 1913;
Best Local Similarity 47.5%; Pred. No. 0.32;
Matches 104; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 146 ccgagattgacgagcagcaagattcaagatgatctgtgaactgtctcctgcaaac 205
DB 233 CCGGGCTGGGAGGCTCCGCCACCTGCTGCTGCCAGTGGCTTCCAGAGATC 174
QY 206 agtgcgagacctgcatgagttgtccaagaatgtgcttcgtatggggaagatgac 265
DB 173 CCTGTGGCCCTGTCAGAGAGTCCCTCTACATCTAGTGGGCGAGCTGAGAGGACTG 114
QY 266 agtgtgtccctcagcagccgaccggttcaagaaagactgtgggtttccaaagtgtgaagc 335
DB 113 GTGGAACCTGCTGAGGAGGAGGGGGGGAAGGCCCAAGGCCCGAGGTCCCGGC 54
QY 326 catgtcgagactgtgcctgtgtgaacgcgttcagaagg 364
DB 53 GTGGGGGCTCTGGGGAGTGAACACCCGCTGGGGGG 15

```

```

WP Sequence split into 6 fragments LOCUS V30458 Accession V30458
WP Fragment Name Begin End
WP V30458_0 1 110000
WP V30458_1 100001 210000
WP V30458_2 200001 310000
WP V30458_3 300001 410000
WP V30458_4 400001 510000
WP V30458_5 500001 534720

Query Match 3.8%; Score 34; DB 1; Length 110000;
Best Local Similarity 57.5%; Pred. No. 4.5;
Matches 61; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 502 accagcaaggtgaacctgtgtagatctctccaccgtctcagccctccgagcagcg 561
DB 31583 ACCAGCAGGTGACCCACCCAGAGTCCGCTCCGCTCGAAGAGCGCTGCACCTGCC 31642
QY 562 gtggtgcgcgtcatctgcatgtcgtgcacagtgctgcgct 607
DB 31643 TGGCATCGCGCATGTCCATGGCCATGGCCAGCATTTGCCCGCTT 31688

RESULT 8
Q94328
ID 094328 standard; cDNA; 3013 BP.
AC 094328:
DT 16-MAY-1996 (first entry)
DE Rabbit telencephalin cDNA.
KW Japanese white rabbit; brain; telencephalon; telencephalin;
KW myotetolagus cuniculus; learning; memory; sensory disorder; ds.
OS Oryctolagus cuniculus.
FH Key Location/Qualifiers
FT cds 125..2863
FT /tag= a
FT /transl_except= pos:1124..1126, aa:Val
FT /transl_except= pos:1814..1816, aa:Thr
FT signal_peptide 125..211
FT /tag= b
FT mat_peptide 212..2860
FT /tag= c
FT /product= telencephalin

PN J07242697-A.
PD 19-SEP-1995.
PF 02-MAR-1994; 032328.
PR 02-MAR-1994; JP-032328.
PA (OSAB-) ZH OSAKA BIOSCIENCE KENKYUSHO.
DR WPI: 95-355283/46.
DR P-PSDB: R75206.
PT DNA encoding telencephalin, a telencephalon specific membranous
PT protein - used in treatment of learning, memory and sensory
PT disorders.
PS Claim 6; Fig 2 to 4; 18pp; Japanese.
CC This sequence encodes a membrane-specific protein isolated from the
CC telencephalon of Japanese white rabbits. The protein, designated
CC telencephalin, is useful for treating diseases associated with the
CC telencephalon (e.g. learning, memory, sensory, emotional and
CC language disorders).
SQ Sequence 3013 BP; 452 A; 1032 C; 1073 G; 456 T;

Query Match 3.8%; Score 33.8; DB 1; Length 3013;
Best Local Similarity 52.5%; Pred. No. 0.95;
Matches 74; Conservative 0; Mismatches 67; Indels 0; Gaps 0;

QY 478 cctccctccctagcaaccacatgctacacgaagtggaacctgtgaagatctccacc 537
DB 1982 CCTAGTCCAGGAGACCCAGCGTCCCTTAGAGACCTGGACCCGGGATCTACTCTGCAC 2041
QY 538 gtctcagccctcgggacacgagtggtgctgcgtcatctgcatgtctggtccaggtg 597
DB 2042 GGCACCAACCCGCTGGGCTCGGGGTCACAAACAGTGTGTAGGCGGAGTGGCGCGC 2101

```

OY 598 ctgcgcctcctcctcctgt 618
 Db 2102 CAGATGACGACTCCACGCT 2122

RESULT 9

N70218
 ID N70218 standard; DNA; 2810 BP.
 AC N70218;
 DE 25-APR-1991 (first entry)
 LEU2 gene of Y.lipolytica.
 KM Leu2: beta-isopropylmalate dehydrogenase; ss.
 OS Yarrowia lipolytica.
 FH Key Location/Qualifiers
 FT cds /tag= a
 FT /product= beta-isopropylmalate dehydrogenase
 FT caat_signal 675..678
 FT /tag= b
 FT tata_signal 721..728
 FT /tag= c
 FT terminator 2088..2137
 FT /tag= d
 FT /note= "homologous to TAG, TA(T)GT, TTT sequence
 of S. cerevisiae proposed by Zaret et al.
 Cell 28, 563(1982) as important for
 transcription termination"

EP-220864-A.
 PD 06-MAY-1987.
 PE 10-OCT-1986: 307839.
 PR 18-OCT-1985: US-789206.
 PA 18-MAR-1986; US-841121.
 PI (PFIZ) PFIZER INC.
 PI Davidow LS, Franke AE, Dezeuw JR;
 DR WPI: 87-124409/18.
 PT New Yarrowia lipolytica transformants - used for expression and
 PT secretion of heterologous proteins; esp. protenimin, and human
 PT anaphylatoxin C5a
 PS Disclosure: Fig 12; 45pp; English.
 CC The Leu2 gene was cloned in pLD25 (EP-138508). It was sequenced and
 CC the correct 194 dehydrogenase coding region was located by
 CC comparison with the known leu2 sequence of S. cerevisiae. The
 CC promoter and terminator sequences are useful for expression of
 CC heterologous proteins.
 CC See also N70213-7, N71339, N71340, N71343-8.
 SQ Sequence 2810 BP; 679 A; 709 C; 701 G; 721 T;

Query Match 3.7%; Score 33; DB 1; Length 2810;
 Best Local Similarity 57.1%; Pred. No. 1.6;
 Matches 60; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

OY 503 ccagcaagtgaaacctgtgaagatctctccacgctccacccctcgagacagcg 562
 Db 1538 CCCCAGATGAATGGATATCATCATCCACCAACATGTTGGCGATATATCTCCGACG 1597
 OY 563 tggctgcgcgtcatctgcagtgctctgcggccaggtgctgctcgct 607
 Db 1598 AGGCTTCGTCATCCCGGTTCTGTGGGTGCTGCTGCCCTCCGCT 1642

RESULT 10

O83184
 ID O83184 standard; DNA; 2810 BP.
 AC O83184;
 DT 03-NOV-1995 (first entry)
 DE Yarrowia lipolytica LEU2 gene.
 KM Leu2 gene; Yarrowia lipolytica; modified; promoter region; expression;
 OS heterologous proteins; Al4trp proinsulin; insulinotropin; ds.
 PA Yarrowia lipolytica.
 FH Key Location/Qualifiers
 FT misc_feature 799
 FT /tag= a

FT FT /note= "start of structural gene"
 FT promoter 693..798
 FT /tag= b
 FT /note= "region is claimed"
 FT promoter 718..798
 FT /tag= c
 FT /note= "promoter region may be modified as shown
 below (claimed)"
 FT promoter 745..798
 FT /tag= d
 FT /note= "region is claimed"
 FT misc_feature 724
 FT /tag= e
 FT /note= "A is changed to G"
 FT misc_feature 725
 FT /tag= f
 FT /note= "T is changed to G"
 FT misc_feature 722
 FT /tag= g
 FT /note= "A is changed to G"
 FT W09506739-A.
 PD 09-MAR-1995.
 PE 30-MAY-1994: IB0128.
 PR 02-SEP-1993; US-117375.
 PA (PFIZ) PFIZER INC.
 PI James IC, Strick CA;
 DR WPI: 95-115454/15.
 PT Modified Yarrowia lipolytica LEU2 gene promoters and genes - also
 PT expression vectors contg them for multiple integration into Y
 PT lipolytica to express and secrete heterologous protein.
 PS Disclosure: Page 25-26; 57pp; English.
 CC The Leu2 gene from Yarrowia lipolytica (shown here) can be modified
 CC within the promoter region and used to express heterologous proteins,
 CC esp. Al4trp proinsulin and insulinotropin. (Also see O83185-200).
 SQ Sequence 2810 BP; 678 A; 708 C; 702 G; 722 T;

Query Match 3.7%; Score 33; DB 1; Length 2810;
 Best Local Similarity 57.1%; Pred. No. 1.6;
 Matches 60; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

OY 503 ccagcaagtgaaacctgtgaagatctctccacgctccacccctcgagacagcg 562
 Db 1538 CCCCAGATGAATGGATATCATCATCCACCAACATGTTGGCGATATATCTCCGACG 1597
 OY 563 tggctgcgcgtcatctgcagtgctctgcggccaggtgctgctcgct 607
 Db 1598 AGGCTTCGTCATCCCGGTTCTGTGGGTGCTGCTGCCCTCCGCT 1642

RESULT 11

V62933/C
 ID V62933 standard; cDNA to mRNA; 5822 BP.
 AC V62933;
 DT 15-JAN-1999 (first entry)
 DE Human mdia Rho targeting protein cDNA.
 KM Rho Protein; mdia; mammalian diaphanous; target protein; human;
 KM Rho protein-combining; proline-combining; chromosome 5q31.2; disease;
 KM respiratory tract; overensitiveness; bronchial asthma; marrow leukemia;
 KM marrow dysplasia syndrome; ss.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 28..3975
 FT /tag= a
 FT /product= "mdia"
 PN J10262680-A.
 PD 06-OCT-1998.
 PE 25-MAR-1997; 090170.
 PR 25-MAR-1997; JP-090170.
 PA (KIRI) KIRIN BREWERY KK.
 DR WPI: 99-002481/01.
 DR P-PSDB; W76734.
 PT Human Rho target protein and its gene - useful for elucidation of

PT mechanisms of respiratory tract disease
 PS Claim 5; Page 31-37; 54pp; Japanese.
 CC This sequence encodes a human Rho target protein, mdia (mammalian
 CC diaphanous). This protein has active type Rho protein-combining ability,
 CC has proline-combining ability, has a M.W. of 150 kDa measured by sodium
 CC dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and its
 CC gene is positioned at human chromosome 5q31.2. The nucleic acid can be
 CC used for the recombinant production of the protein. The protein is useful
 CC for the elucidation of mechanism of diseases such as respiratory tract
 CC overensitiveness, bronchial asthma, acute marrow leukaemia and marrow
 CC dysplasia syndrome.
 SQ Sequence 5822 BP; 1468 A; 1448 C; 1504 G; 1402 T;

Query Match 3.5%; Score 31; DB 1; Length 5822;
 Best Local Similarity 56.3%; Pred. No. 9.6;
 Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 550 cggagacacggcggtggtccgtcatctgcaagtctctggcacaagggtgctgcgtcgc 609
 DB 4203 CTGGCTAAGTCGGGCTCAGGCTCCCTCGACACTGCCCTTCTCTGCGCTCCAGGACG 4144
 QY 610 tcatcctgtgtcatctactgcaaggagcagttcatgagaa 652
 DB 4143 TGAAGCTGTATGTGATGTTGAGAGACGACGCGCCAGAGAA 4101

RESULT 12
 V62934
 ID V62934 standard; cDNA to mRNA; 5822 BP.
 AC V62934;
 DT 15-JAN-1999 (first entry)
 DE Human mdia Rho targeting protein cDNA #2.
 KM Rho protein: mdia; mammalian diaphanous; target protein; human;
 KM Rho protein-combining; proline-combining; chromosome 5q31.2; disease;
 KM respiratory tract; overensitiveness; bronchial asthma; marrow leukaemia;
 KM marrow dysplasia syndrome; ss.
 OS Homo sapiens.
 PN J1026280-A.
 PD 06-OCT-1998.
 PF 25-MAR-1997; 090170.
 PR 25-MAR-1997; JP-090170.
 PA (KIRI) KIRIN BREWERY KK.
 DR WPI; 99-002481/01.

PT Human Rho target protein and its gene - useful for elucidation of
 PT mechanisms of respiratory tract disease
 PS Disclosure; Page 37-39; 54pp; Japanese.
 CC This sequence encodes a human Rho target protein, mdia (mammalian
 CC diaphanous). This protein has active type Rho protein-combining ability,
 CC has proline-combining ability, has a M.W. of 150 kDa measured by sodium
 CC dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and its
 CC gene is positioned at human chromosome 5q31.2. The nucleic acid can be
 CC used for the recombinant production of the protein. The protein is useful
 CC for the elucidation of mechanism of diseases such as respiratory tract
 CC overensitiveness, bronchial asthma, acute marrow leukaemia and marrow
 CC dysplasia syndrome.
 SQ Sequence 5822 BP; 1402 A; 1504 C; 1448 G; 1468 T;

Query Match 3.5%; Score 31; DB 1; Length 5822;
 Best Local Similarity 56.3%; Pred. No. 9.6;
 Matches 58; Conservative 0; Mismatches 45; Indels 0; Gaps 0;

QY 550 cggagacacggcggtggtccgtcatctgcaagtctctggcacaagggtgctgcgtcgc 609
 DB 1620 CTGGCTAAGTCGGGCTCAGGCTCCCTCGACACTGCCCTTCTCTGCGCTCCAGGACG 1679
 QY 610 tcatcctgtgtcatctactgcaaggagcagttcatgagaa 652
 DB 1680 TGAAGCTGTATGTGATGTTGAGAGACGACGCGCCAGAGAA 1722

RESULT 13

027167
 ID 027167 standard; DNA; 3471 BP.
 AC 027167;
 DT 27-JAN-1993 (first entry)
 DE Delta-endotoxin gene.
 KM B.t.PS50C; Colorado potato beetle; Leptinotarsa decemlineata;
 KM Pseudomonas; expression vector; ss.
 OS Bacillus thuringiensis.
 PN EP-48537-A.
 PD 12-AUG-1992.
 PF 16-JAN-1992; 300366.
 PR 16-JAN-1991; US-642112.
 PA (MYCO) MYCOGEN CORP.
 PI Roncerada L, Payne JM, Sick AJ;
 DR WPI; 92-270497/33.
 P-PDB; R23997.

PT New insecticidal Bacillus thuringiensis strain NRRL B-18746 - and
 PT toxin, DNA encoding it, vectors and transformed cells, effective
 PT against Coleoptera
 PS Claim 8; Page 9-11; 16pp; English.
 CC The sequence given encodes an approx. 130 kD protien known as Bacillus
 CC thuringiensis PS50C (B.t.PS50C). This novel delta-endotoxin has been
 CC shown to be active against the Colorado potato beetle (Leptinotarsa
 CC decemlineata). This toxin gene can be transformed into suitable hosts
 CC such as Pseudomonas, which can then be applied to the environment of
 CC coleopteran insects where they will proliferate and be ingested by the
 CC insects. This sequence can be inserted into an expression vector which
 CC contains a promoter/operator region, a ribosome binding site,
 CC polyadenylation signals, etc. This will allow transcription and
 CC translation of the cells in the appropriate host.
 SQ Sequence 3471 BP; 1200 A; 567 C; 751 G; 953 T;

Query Match 3.4%; Score 30.6; DB 1; Length 3471;
 Best Local Similarity 51.9%; Pred. No. 10;
 Matches 69; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 757 tagagcstaacagcctactagaatacaagatcaggaagaacagcctctcaggaatc 816
 DB 1628 TAGTACAGGGTCTCGATTACAGAGAGATATATTAATAAACCAATCTTAGATAT 1687
 QY 817 caggcctcctaggagatcgtgcaagcgtgtgattctcgaagctaccaggaataata 876
 DB 1688 TAGGACCTTTGGCGTTACAGTAATAGCGTTCATTACAAAGATATCGTTAAGATTTC 1747

QY 877 aagltgtctatc 889
 DB 1748 GCTATGCGCTCTAC 1760

RESULT 14
 Q28940
 ID Q28940 standard; DNA; 3471 BP.
 AC Q28940;
 DT 24-FEB-1993 (first entry)
 DE B.thuringiensis PS50C insecticidal protein gene.
 KM B.t.PS50C; NRRL B-18746; PMTC1638; Colorado potato beetle;
 KM coleopteran-active toxin; ss.
 OS Bacillus thuringiensis.
 FH key location/Qualifiers
 FT mat_peptide 1..3471
 FT FT
 FT /+tag= a
 FT /product= toxin
 FT /note= "active against Coleoptera"

CA2059242-A.
 17-JUL-1992.
 PF 13-JAN-1992; 059242.
 PR 16-JAN-1991; US-642112.
 PR 02-JAN-1992; US-812180.
 PA (MYCO) MYCOGEN CORP.
 PI Roncerada L, Payne JM, Sick AJ;
 DR WPI; 92-323784/40.
 P-PDB; R27343.

PT New *Bacillus thuringiensis* isolate - used for preparing toxin,
PT DNA and transformed hosts for controlling coleoptera insect pests
PS Claim 13; Page 18-20; 33pp; English.
CC Total cellular DNA was prep'd. from B.t. p550C. The DNA was partially
CC digested with *Sma*I and 9-23kb fragments were ligated into BamH-
CC digested lambda GEM-11. Packaged phage were plated on E.coli KW251
CC cells and screened using a radiolabelled probe. Positive plaques
CC were purified and rescreened. Single isolated purified plaques were
CC used to infect E.coli KW251 in liquid culture. DNA was isolated,
CC digested with *Xho*I and large fragments ligated to *Xho*I-cut pTf1uett.
CC The ligation mixt. was introduced into competent E.coli NM522 cells.
CC Plasmid pMYC1638 contained a 12kb *Xho*I insert from which the 3471bp
CC sequence of the toxin gene was determined.
50 Sequence 3471 bp; 1202 A; 570 C; 745 G; 954 T;

This Page Blank (uspto)

| | | | | | | | | | |
|---|----|------|-----|-----|----|------------|------------|------------|------------|
| C | 3 | 67.6 | 7.1 | 404 | 20 | AA495217 | AA495217 | fA04d10 | AA495217 |
| C | 4 | 36.4 | 4.6 | 404 | 22 | AA552514 | AA552514 | nK12n08.s | AA552514 |
| C | 5 | 36 | 4.0 | 399 | 29 | A1084589 | A1084589 | OX63609.s | A1084589 |
| C | 6 | 35.2 | 3.9 | 552 | 23 | A1063417 | A1063417 | Gt03215.S | A1063417 |
| C | 7 | 34.8 | 3.9 | 416 | 25 | AA692849 | AA692849 | vT57h10.s | AA692849 |
| C | 8 | 34.6 | 3.9 | 301 | 26 | AA743377 | AA743377 | nv18e11.s | AA743377 |
| C | 9 | 34.4 | 3.9 | 386 | 12 | H92111 | H92111 | yS84f03.r1 | H92111 |
| C | 10 | 34.4 | 3.9 | 275 | 27 | AA829315 | AA829315 | oD05f09.s | AA829315 |
| C | 11 | 34.2 | 3.8 | 295 | 17 | AA359174 | AA359174 | EST68148 | AA359174 |
| C | 12 | 34.2 | 3.8 | 433 | 21 | AA279548 | AA279548 | zS86B08.r1 | AA279548 |
| C | 13 | 34.2 | 3.8 | 434 | 30 | A1131315 | A1131315 | uP90e02.r1 | A1131315 |
| C | 14 | 34 | 3.8 | 383 | 14 | AA055710 | AA055710 | z175e03.r1 | AA055710 |
| C | 15 | 33.8 | 3.8 | 553 | 31 | A1266650 | A1266650 | qM45d03.x | A1266650 |
| C | 16 | 33.6 | 3.8 | 467 | 10 | R02555 | R02555 | yE80b07.s1 | yE80b07.s1 |
| C | 17 | 33.6 | 3.8 | 585 | 13 | W02452 | W02452 | zA47B01.r1 | W02452 |
| C | 18 | 33.4 | 3.7 | 421 | 11 | H06838 | H06838 | Y184b02.r1 | H06838 |
| C | 19 | 33.4 | 3.7 | 466 | 11 | H14764 | H14764 | Ym24h03.r1 | H14764 |
| C | 20 | 33.4 | 3.7 | 497 | 13 | N90576 | N90576 | zB42C02.r1 | N90576 |
| C | 21 | 33.4 | 3.7 | 309 | 14 | W35818 | W35818 | mC14g03.r1 | W35818 |
| C | 22 | 33.4 | 3.7 | 272 | 18 | AA341295 | AA341295 | EST46907 | AA341295 |
| C | 23 | 33 | 3.7 | 487 | 11 | H20423 | H20423 | yN56c12.r1 | H20423 |
| C | 24 | 33 | 3.7 | 589 | 13 | W37947 | W37947 | zC13d03.r1 | W37947 |
| C | 25 | 33 | 3.7 | 504 | 22 | W10615 | W10615 | ma39d10.r1 | W10615 |
| C | 26 | 33 | 3.7 | 233 | 29 | R1CC10050A | R1CC10050A | Rice cDNA, | D21918 |
| C | 27 | 33 | 3.7 | 333 | 30 | A1156639 | A1156639 | ue53p03.r1 | A1156639 |
| C | 28 | 32.6 | 3.7 | 427 | 11 | T86351 | T86351 | yD84f11.s1 | T86351 |
| C | 29 | 32.6 | 3.7 | 506 | 11 | H14236 | H14236 | yM63a10.s1 | H14236 |
| C | 30 | 32.6 | 3.7 | 415 | 11 | R63653 | R63653 | yM10D03.s1 | R63653 |
| C | 31 | 32.4 | 3.6 | 450 | 11 | R91903 | R91903 | yQ04h10.r1 | R91903 |
| C | 32 | 32.4 | 3.6 | 474 | 27 | AA845505 | AA845505 | zK04B11.s | AA845505 |
| C | 33 | 32.2 | 3.6 | 437 | 21 | AA292015 | AA292015 | zC28C10.s | AA292015 |
| C | 34 | 32.2 | 3.6 | 477 | 24 | AA061232 | AA061232 | zU81h04.s | AA061232 |
| C | 35 | 32.2 | 3.6 | 435 | 28 | A1007948 | A1007948 | EST2023939 | EST2023939 |
| C | 36 | 32.2 | 3.6 | 357 | 28 | A1013679 | A1013679 | EST208355 | EST208355 |
| C | 37 | 32.2 | 3.6 | 339 | 30 | A1100988 | A1100988 | EST2102777 | EST2102777 |
| C | 38 | 32.2 | 3.6 | 402 | 30 | A1140948 | A1140948 | UT-R-BT0 | A1140948 |
| C | 39 | 32 | 3.6 | 413 | 13 | W76336 | W76336 | zD60G01.r1 | W76336 |
| C | 40 | 32 | 3.6 | 382 | 29 | A1037602 | A1037602 | uB62g10.r1 | A1037602 |
| C | 41 | 31.8 | 3.6 | 440 | 15 | AA165941 | AA165941 | mS50e12.r1 | AA165941 |
| C | 42 | 31.8 | 3.6 | 512 | 30 | A11819167 | A11819167 | ue23f10.r1 | A11819167 |
| C | 43 | 31.6 | 3.5 | 368 | 27 | AA831737 | AA831737 | oC93a12.s | AA831737 |
| C | 44 | 31.6 | 3.5 | 428 | 30 | A1155141 | A1155141 | uD59f11.r1 | A1155141 |
| C | 45 | 31.4 | 3.5 | 359 | 10 | T65828 | T65828 | yC11c06.r1 | T65828 |

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:261704

Seq primer: ETPRimer
 High quality sequence stop: 345.
 Location/Qualifiers

FEATURES

source

1..401
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TTTTACCACTGTAAGTGGAGCGGCGGGAATTTTCTTTTCTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2 1] double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of the modified
 pT73 vector. Library went through one round of
 normalization, and was constructed by Bento Soares and
 M.Fatima Bonaldo."
 /db_xref="taxon:10090"
 /clone="427152"
 /clone_1ib="Soares mouse embryo NbME13.5 14.5"
 /sex="unknown"
 /tissue_type="embryo"
 /dev_stage="13.5-14.5dpc total fetus"
 /lab_host="DH10B"
 <1..>3401
 92 a 96 c 128 g 85 t

BASE COUNT
 MRNA
 ORIGIN

Query Match 37.3%; Score 333.4; DB 13; Length 401;

Best Local Similarity 98.3%; Pred. No. 5e-93;
 Matches 337; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 19 gacaaagggaagaaataaacagcttggtgagagccatggcactcaaggtcctactcta 78.
 |||
 DB 59 GCGGAGTAGCAGGAATAAACGTTGTGTGAGAGCATGCGACTCAAGGCTTACCTCTA 118
 QY 79 caaagagaggtgtcttcgttcgtccattcttcttactccactgtgcattgaagtgagt 138
 |||
 DB 119 CACAGGACGCTCTTCGCTGCGCATCTCTCTACTCACCCTGCATGATTAAGTGAGT 178
 QY 139 tgcgaaccggaattgcagagcagcaggaattcaagatgcattggaactgttcctc 198
 |||
 DB 179 TCGGAACCGGAATTCAGAGCAGCAGGAATTCAGGATTCATGAATCTGTCTCTC 238
 QY 199 tgcgaacagtgccagactgtgcattgttccaaggaatgtgtccttcggtccta 258
 |||
 DB 239 TCGAAACAGTGCAGACTGCGCATGAGATTGCCAAGGAATGTGGCTTCGGCTTGGGGAG 298
 QY 259 gatgcacagtggtgtccctccagcagccgacacgggttcaaggaagactgggggttc 318
 |||
 DB 299 GATGACACAGTGTGCTTCGCAAGCGCAGCCGGTTCAAGGAACATGGGGTTCCAGAA 358
 QY 319 tctgaagccatgtgcagactgtgcctcgtgtgaacgcgttcaga 361
 |||
 DB 359 TGTAAACCATGTGCGGACTGTGCGCTGTGTAACCGCTGTCTAGA 401

RESULT 2
 AA036247 358 bp MRNA EST 26-AUG-1996
 LOCUS mi74a03.r1 Soares mouse p3NMF19.5 Mus musculus cDNA clone 472300
 DEFINITION 5', mRNA sequence.
 ACCESSION AA036247
 NID 91509376

KEYWORDS

SOURCE

ORGANISM

REFERENCE

EST.
 house mouse.
 Mus musculus
 Eukaryote; mitochondrial eukaryotes; Metazoa; Chordata;
 Vertebrata; Eutheria; Rodentia; Sciurognathi; Myomorpha; Muridae;
 Murinae; Mus.
 1 (bases 1 to 358)
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)

TITLE

JOURNAL

COMMENT

Contact: Maria M/Mouse EST Project
 WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@watson.wustl.edu
 This clone is available royalty-free through LNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:283044
 Seq primer: -28M13 rev2 from Amersham
 High quality sequence stop: 347.
 Location/Qualifiers

FEATURES

source

1..358
 /organism="Mus musculus"
 /note="Vector: pT73D (Pharmacia) with a modified
 polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
 was primed with a Not I - oligo(dT) primer [5'
 TTTTACCACTGTAAGTGGAGCGGCGGGAATTTTCTTTTCTTTT
 T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
 14.5dpc embryos [total RNA provided by Minoru Ko, Wayne
 State Univ., from 2 1] double-stranded cDNA was ligated to
 Eco RI adaptors (Pharmacia), digested with Not I and
 cloned into the Not I and Eco RI sites of a modified pT73 vector
 (Pharmacia). Library went through one round of
 normalization to a Cot = 5. Library constructed by Bento
 Soares and M.Fatima Bonaldo. RNA was kindly provided by
 Dr. Minoru Ko (Wayne State University)."
 /db_xref="taxon:10090"
 /clone="472300"
 /clone_1ib="Soares mouse p3NMF19.5"
 /dev_stage="19.5 dpc total fetus"
 /lab_host="DH10B (ampicillin resistant)"
 <1..>358
 81 a 98 c 110 g 69 t

BASE COUNT
 MRNA
 ORIGIN

Query Match 35.9%; Score 321; DB 14; Length 358;
 Best Local Similarity 100.0%; Pred. No. 3.4e-89;
 Matches 321; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tctgggcccgggggggaagcaaggggaagaaataaacagcttggtgagagccatgga 60
 |||
 DB 38 TCCGGGCGCGCGGCGGAGCAAGGGAAGGAATAAACGTTGTGTGAGAGCATTGCGCA 97
 QY 61 ctcaaggtcctaccttaacaggaagcagtggtcttcgtgcattcttcttactccac 120
 |||
 DB 98 CTCAGAGTCTTACTTACACAGAGAGGAGTGTCTTTCGCTCCATTTCTTCTACTCCAC 157
 QY 121 ctggcactgtaagtgtgcttcgcaaacacggagattcgaggcagcaggaaattcaagatcga 180
 |||
 DB 158 CTGGCATGTAAAGTGAAGTTCGGAACCGGAGATGTGAGCAGCAGGAATTCAGAGATCA 217
 QY 181 tctggaaactgtgtctcttcgcaaacagtcggactgtgcattgagtggttccaaggaatgt 240
 |||
 DB 218 TCTGGAATCTGTCTCTGCAACAGTCCGAGCTGTGAGATGAGATTGTCCAAAGGATGT 277
 QY 241 ggcttcgctatggggagagatgcacagtggtgtcccttcgagggcgacacgggttcaagga 300
 |||

| | | | |
|-----------------------|---|---|---------------------------------------|
| QY | 631 | gcaagagcgagctcatgtgagagaagaaccacgctgtgaagctccatccctctgtccactg | 690 |
| Db | 210 | GGAGCAGATTTATATATGGAAAAATATTAAAGGAGGAAGATGTGTACTGTGTACACAG | 151 |
| QY | 691 | tgaagtgaagctgtgttagcatgtcaccccaagaagcttctcaagacc | 736 |
| Db | 150 | TGATTACATCCAGGCATTATTGCCAGCCAGATTTATAAATATGCC | 105 |
| RESULT | 5 | | |
| A1084589 | A1084589 | 399 bp | mRNA |
| LOCUS | ox3e09.s1 | Soares_NHMPU.s1 | Homo sapiens cDNA clone IMAGE:1661032 |
| DEFINITION | 3', mRNA sequence. | | |
| ACCESSION | A1084589 | | |
| NID | 93423012 | | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; | | |
| AUTHORS | Primates; Catarrhini; Homnidae; Homo. | | |
| TITLE | 1 (bases 1 to 399) | | |
| JOURNAL | NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap . | | |
| COMMENT | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index | | |
| FEATURES | Seq primer: -40m13 fwd. ET from Amersham. | | |
| SOURCE | Location/Qualifiers | | |
| | 1..399 | | |
| | /organism="Homo sapiens" | | |
| | /note="Organ: mixed (see below); Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI. Equal amounts of plasmid DNA from three normalized libraries (melanocyte 2NBH, pregnant uterus NBHPU, and fetal heart NBH19W) were mixed, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 3 libraries. The pools consisted of 1 M.A.G.E. clones 260232-265223, 340488-345479, and 484488-489479." | | |
| | /db_xref="taxon:9606" | | |
| | /clone="IMAGE:1661032" | | |
| | /clone_1lb="Soares_NHMPU.s1" | | |
| | /tissue_type="Pooled human melanocyte, fetal heart, and pregnant uterus" | | |
| | /lab_host="DH10B" | | |
| BASE COUNT | 112 a | 92 c | 77 g |
| ORIGIN | | | |
| Query Match | 4.0%; | Score 36; | DB 29; |
| Best Local Similarity | 51.9%; | Pred. No. 0.76; | Length 399; |
| Matches | 81; Conservative | 0; Mismatches | 75; Indels |
| QY | 613 | tcctctgtgtcactactcaagaagcgagctatgtgagaagaaccacgctgtaagctcc | 672 |
| Db | 6 | TCTCTGGAATATATCTCATATTAAGAAAAAGCTTTTAAGAGAGAAAAACCCACACACT | 65 |
| QY | 673 | catccctctgtctcaactgtgaagtgagctgtgttaagcatgttaccocaaagctccaaga | 732 |
| Db | 66 | CAGGCTCTTATATCCCTCCCTGCTGTTCTTTTCAGAGTTCTGACCTCAAGATTCTGAAAA | 125 |
| QY | 733 | caactgtgtcgaagactaagacctttagagacatcaaac | 768 |
| Db | 126 | GTACAGAGTCTACCTTCATCCCTTAGCAAGAGAAAC | 161 |

| RESULT | 6 | 552 bp | mRNA | EST | 24-NOV-1998 |
|-----------------------|--|---|----------------|-------------|-------------|
| LOCUS | AI063417 | | | | |
| DEFINITION | GH03215.5prime GH Drosophila melanogaster head pot2 Drosophila melanogaster cDNA GH03215 5prime similar to U23545: HK Fgmn0001203 PID:902000 SPTREMBL:024052, mRNA sequence. | | | | |
| ACCESSION | AI063417 | | | | |
| KEYWORDS | EST. | | | | |
| KEYWORDS | fruit fly. | | | | |
| SOURCE | Drosophila melanogaster | | | | |
| ORGANISM | Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. | | | | |
| REFERENCE | 1 (bases 1 to 552) | | | | |
| AUTHORS | Harvey,D., Hong,L., Evans-Holm,M., Pendleton,J., Su,C., Brokstein,P., Lewis,S. and Rubin,G.M. | | | | |
| TITLE | BDGP/HMI Drosophila EST Project | | | | |
| JOURNAL | Unpublished (1997) | | | | |
| COMMENT | Contact: Harvey, D. G. M. Rubin-Molecular and Cell Biology University of California Berkeley 539 LSA, Berkeley, CA 94720-3200, USA Fax: 510 643 9947 Email: http://www.fruitfly.org/EST , est@fruitfly.berkeley.edu Plate: 32 row: B column: 3 High quality sequence stop: 540. Location/Qualifiers 1..552 /organism="Drosophila melanogaster" /note="Organ: head; Vector: pot2; Site_1: EcoRI; Site_2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library." /db_xref="taxon:7227" /clone="GH03215" /clone_1ib="GH Drosophila melanogaster head pot2" /sex="male and female" /dev_stage="adult" /lab_host="DHS - alpha" 114 a 168 c 169 g 101 t | | | | |
| BASE COUNT | 114 a | 168 c | 169 g | 101 t | |
| ORIGIN | | | | | |
| Query Match | 3.9% | Score 35.2; | DB 23; | Length 552; | |
| Best Local Similarity | 50.0%; | Pred. No. 1.5; | | | |
| Matches | 88; | Conservative 0; | Mismatches 88; | Indels 0; | Gaps 0 |
| Oy | 100 | gccattctcttctactccacctgcgcatctaaagtgagttgcgaacgcgagatgcagg | 159 | | |
| | | | | | |
| Db | 6 | gccattgacgtacctgattacgacgacggcctgggcaaatgatttggggcaccgcctcatgagac | 65 | | |
| Oy | 160 | cagcagaattcaagaagatcgatctgtaaacctgtatcccttgcaaacagtcgcgaacctggc | 219 | | |
| | | | | | |
| Db | 66 | cagctggacatctatcgacggcccttaccacactcccccagttcaactgcattcaccccat | 125 | | |
| Oy | 220 | atgagatgtcccaagaagatgtgcttcgcgtatctggggggagatgcacagatgtgtgc | 275 | | |
| | | | | | |
| Db | 126 | gttgagcagtcacacgattacattatgtctgtccgcaaaagtcgacactctacctgcc | 181 | | |
| RESULT | 7 | 416 bp | mRNA | EST | 16-DEC-1997 |
| LOCUS | AA692849 | | | | |
| DEFINITION | v57h10.s1 Knowles Solter mouse 2 cell Mus musculus cDNA clone 1124803 5' similar to FR:Q14136 Q14136 KIAA0122 PROTEIN ;, mRNA sequence. | | | | |
| ACCESSION | AA692849 | | | | |
| KEYWORDS | EST. | | | | |
| KEYWORDS | house mouse. | | | | |
| SOURCE | Mus musculus | | | | |

MGI:907486
Trace considered overall poor quality
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 1.

FEATURES

source

Location/Qualifiers

1..434
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: mammary gland; Vector: pT73D-Pac
(pharmacia) with a modified polylinker. Site.1: Not I;
Site.2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer [5'
TGTTCACATCTGAAGTGGACCGCCGGAATGTTTGTGTGTGTGTGTGTGT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru KO, Wayne State Univ. Library
constructed and normalized by Bento Soares and M. Fatima
Bonaldo.
/db_xref="taxon:10090"
/clone="1395770"
/clone.lib="Soares mouse mammary gland NbMG"
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"

BASE COUNT

121 a 84 c 131 g 98 t

ORIGIN

Query Match 3.8%; Score 34.2; DB 30; Length 434;
Best Local Similarity 60.0%; Pred. No. 2.8;
Matches 57; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

QY 254 ggagagatgacagctgtgtgctcctgcagcgccgacgggtccaaggaactgtgttc 313
|||||
Db 133 ggagagagccatccgggtcttggagcagagagacaaagtaagattgaattaa 192
|||||

QY 314 agaagctaaagcattgtgcgactgtgcgtgtgtg 348
|||||
Db 193 agaatataaaccttctgttgggtgttggaagtgc 227
|||||

RESULT 14

AA055710 383 bp mRNA EST 17-SEP-1996
LOCUS 217se03.r1 Stratagene colon (#937204) Homo sapiens cDNA clone
DEFINITION 510460 5' similar to contains 11.B5 11 repetitive element ; , mRNA
sequence.
ACCESSION AA055710
NID g1548048
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 383)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maira,M.,
Paterson,J., Rifkin,L., Rohlfing,T., Tan,F., Trevasakis,E.,
Waterston,R., Williamson,A., Wohldmann,P. and Wilson,R.
WashU-Merck EST Project
Unpublished (1995)

TITLE

JOURNAL

COMMENT

Contact: Wilson RK
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.

Seq primer: -28m13 rev2 from Amersham
High quality sequence stop: 260.

FEATURES

source

Location/Qualifiers

1..383
/organism="Homo sapiens"
/note="Organ: colon; Vector: pBluescript SK-; Site.1:
EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. T-84 colonic epithelial cell line. Average
insert size: 1.0 kb; Uni-ZAP XR vector; -5' adaptor
sequence: 5' GAATTCGGCAG 3' -3' adaptor sequence: 5'
CTCGAGTGTGTGTGTGTGTGT 3'."
/db_xref="taxon:9606"
/clone="510460"
/clone.lib="Stratagene colon (#937204)"
/lab_host="SOLR cells (kanamycin resistant)"

BASE COUNT

120 a 73 c 73 g 72 t 45 others

ORIGIN

Query Match 3.8%; Score 34; DB 14; Length 383;
Best Local Similarity 51.4%; Pred. No. 3.1;
Matches 76; Conservative 0; Mismatches 72; Indels 0; Gaps 0;

QY 741 tggacctaaagaccttaagacatcacacagcacttaagataaagatgcaggaaga 800
|||||
Db 149 TTTAACTTAAAGACCTTCTGACACAGCAAAAGANBACTACATCAGTGAACGACCACTTA 208
|||||

QY 801 ggccttcaggaatctcagagcctcctcaggaatctgcagagctgtgattc 860
|||||
Db 209 CACATGGGAGAAATTTTCGCACTACTCATCTGACAAAGGCTATATCCAGAAATC 268
|||||

Db

269 TACAATGAACCTCAACCAATTTACNAGA 296
|||||

RESULT 15

A1266650 553 bp mRNA EST 16-NOV-1998
LOCUS qm45d03.x1 Soares_placenta_8to9weeks_2NDHP809W Homo sapiens cDNA
DEFINITION clone IMAGE:1891685 3' similar to contains CER.tl CER repetitive
element ; , mRNA sequence.
ACCESSION A1266650
NID g3884808
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 553)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor gene index
JOURNAL Unpublished (1997)
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40up from Gibco
High quality sequence stop: 346.
Location/Qualifiers

FEATURES

source

1..553
/organism="Homo sapiens"
/note="Organ: placenta; Vector: pT73D (pharmacia) with a
modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st
strand cDNA was primed with a Not I - Oligo(dT) primer [5'
TGTTCACATCTGAAGTGGACCGCCGGAATGTTTGTGTGTGTGTGTGT
T 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector


```
(Pharmacia). Library constructed by Bento Soares and  
M.Fatima Bonaldo."  
/db_xref="taxon:9606"  
/clone="IMAGE:1891685"  
/dev_stage="two placenta_8to9weeks_2NDHP8to9w"  
/dev_stage="two placenta: one from 8 weeks and another  
from 9 weeks post conception"  
/lab_host="DH10B (ampicillin resistant)"  
BASE COUNT      130 a      144 c      190 g      89 t  
ORIGIN
```

```
Query Match      3.8%; Score 33.8; DB 31; Length 553;  
Best Local Similarity 47.1%; Pred. No. 4.1; Indels 0; Gaps 0;  
Matches 104; Conservative 0; Mismatches 117; Indels 0; Gaps 0;  
  
QY 395 gcgggggactgctgccaagatttaccggaagacaactggtgttttcaagacatgg 454  
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 200 GGGGGGGCCCCCTGGCAAGCCTGACGAGATGATAGACAGGGCGGACTGCGGCCAGGA 259  
  
QY 455 agtgtgtgcccctgcggaagaccacccctccctacgaaccactgtaccagcaaggtga 514  
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 260 GGAAGGTGGGGTGGGAAGCCCTCGGCCCTCGAGCTGAGGGTGAGACCCAGGCATGTGG 319  
  
QY 515 acctgtgaagatcctctccacogtctccagccctcgggaacgycggtgtgcgtca 574  
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 320 TCCCGCCGACCTGCTGCTGCCAGACGCCGCGGATCCCGGGGCTGCTGTCGCCAAGGCA 379  
  
QY 575 tctgpaagtctctggccaagctgtgctgcctcatcc 615  
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||  
DB 380 GGTGGAAGCTAGCGGTTCGAGGTTTCTGAGACTGAGGCCCTTCC 420
```

Search completed: May 13, 1999, 19:44:13
Job time: 11318 sec

This Page Blank (uspto)

| Accession | U23545 |
|---------------------------|---|
| NID | 9901999 |
| KEYWORDS | |
| SOURCE | |
| ORGANISM | fruit fly. Drosophila melanogaster |
| REFERENCE | Eukaryotes; mitochondrial eukaryotes; Metazoa; Arthropoda; Tracheata; Insecta; Pterygota; Diptera; Brachycera; Muscomorpha; Ephydroloidea; Drosophilidae; Drosophila. |
| AUTHORS | 1 (bases 1 to 546) Chouinard,S.W., Wilson,G.F., Schlimgen,A.K. and Ganetzky,B. |
| TITLE | A potassium channel beta subunit related to the aldo-keto reductase superfamily is encoded by the Drosophila hyperknetic locus |
| JOURNAL | Proc. Natl. Acad. Sci. U.S.A. 92 (15), 6763-6767 (1995) |
| MEDLINE | 93550149 |
| REFERENCE | 2 (bases 1 to 3520) |
| AUTHORS | Chouinard,S.W. |
| TITLE | Direct Submission |
| JOURNAL | Submitted (28-MAR-1995) Scott W. Chouinard, Genetics, University of Wisconsin-Madison, 445 Henry Mall, Madison, WI 53706, USA |
| FEATURES | Location/Qualifiers |
| Source | 1..3520 |
| | /organism="Drosophila melanogaster" |
| | /db_xref="taxon:7227" |
| | /clone="HC206" |
| | /clone_lib="Drosophila head library--Meyerowitz" |
| | /chromosome="X" |
| | /map="9B7-8" |
| | /tissue_type="head" |
| | /dev_stage="adult" |
| gene | 1..3520 |
| | /note="Hyperknetic" |
| | /gene="Hk" |
| | /allele="" |
| | /db_xref="FlyBase:FBgn0001203" |
| | 710..2350 |
| CDS | |
| | /gene="Hk" |
| | /codon_start=1 |
| | /db_xref="FlyBase:FBgn0001203" |
| | /product="Hk protein" |
| | /db_xref="PID:g902000" |
| | /translation="MSMAICNLNGDGSAAOSTSGOSPATAAPLPLPHSHSHLTOP ESTPPLLGHEHSGSAPBEGSGGCVADGATVSIEMPTVADGCVPLPLPPOSTQ PLMALPRLNLFITGPTTQQLMTIGNGAMV1PTPNDNNNNNNNDTSDENPTQ LYRRLAPIASIDMEEFSGRSISGSNPALPLRHSPTPTGLRYKRLKNGSGLRISNG LGTPVSPGVSDQAAILKLAESGINLEFIDEASESETEIGILQACMKRAYIV TTKYVTSPEGEELSRKHITTECVASLORLOIYIVIIHKRDKPCMEYRASY VIOGKAMNCTARMSOVEINEATNRCNPICPIYEOSETHFCEKEKELIPEMY NKIVGICLMANCPMALSDTONGDKLPLPGSEFTKFSWTEDEINNAALSPGWSG KDRIDEGRRHCDRLRLAALAEKLGCSPTOLISMSLKHBPVOCILLAGSABQLHOS LOSLOLPLRSSSYMLELERILENKPVRPPIWISLALR" |
| | 3520 |
| polyA_site | |
| | /gene="Hk" |
| | /note="g A nucleotides; putative" |
| | /db_xref="FlyBase:FBgn0001203" |
| BASE COUNT | 1054 a 873 c 897 g 696 t |
| ORIGIN | |
| Query Match | 5.9%: Score 36.8; DB 3; Length 3520; |
| Best Local Similarity | 47.8%: Pred. NO. 1.2; |
| Matches 107; Conservative | 0; Mismatches 117; Indels 0; Gaps 0; |
| Db | 113 gccatgacacaaagatcctactacacaaagacgagctgcttcgctgcacatctctc 172 |
| | |
| Db | 1652 GTCATCATCCACACAGCGCGGATCCATGTGTCTCCCATGGAGGTGGTCCGCGCATGACGTAC 1711 |
| | |
| Db | 173 ctactcacctgagcattgtaaaagtgaattgctgcgaacacggagatgtgcagcagcaggaattc 232 |
| | |
| Db | 1712 GTGATACAGACAGGGCTGGGCAATGATATTGGGGCACCGCTCATGTGAGCCAGGTGAGATC 1771 |
| | |
| Db | 233 aagatgatctcggaactgtgtctctctgcgaacagctgcgaactcgtgcattgagttgtcc 292 |
| | |
| Db | 1772 ATGGAGGCGCTATACCACTGCGCGAGTTCACATGTCATCAGCCCATGTGTGAGAGCATGCC 1831 |
| | |

```

OY 293 aaggaatgtagctgcgtatcggtggaagatgcacagtgtgctc 336
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1832 GAGTGCATATGTTCTGTCCGAAAAAGTGCAGAACTCTACCTGC 1875

RESULT 4
HS426N21/c
LOCUS
DEFINITION
ACCESSION
NID
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Submitted (06-NOV-1996) Wellcome Trust Genome Campus, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humangenesanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Jan 12, 1997 this sequence version replaced gi:166361.
IMPORTANT: This sequence is unfinished and does not necessarily
represent the correct sequence. Work on the sequence is in progress
and
the release of this data is based on the understanding that the
sequence may change as work continues. The sequence may be
contaminated
with foreign sequence from E.coli, yeast, vector, phage etc. Order
of segments is not known; 800 n's separate segments. Unfinished
sequence: d426N21 Contig_ID: 00151 Length: 2531 bp Unfinished
sequence: d426N21 Contig_ID: 00106 Length: 1888 bp Unfinished
sequence: d426N21 Contig_ID: 00040 Length: 862 bp Unfinished
sequence: d426N21 Contig_ID: 02103 Length: 5017 bp Unfinished
sequence: d426N21 Contig_ID: 02094 Length: 2158 bp Unfinished
sequence: d426N21 Contig_ID: 02176 Length: 4992 bp Unfinished
sequence: d426N21 Contig_ID: 02098 Length: 2049 bp Unfinished
sequence: d426N21 Contig_ID: 02159 Length: 950 bp Unfinished
sequence: d426N21 Contig_ID: 00949 Length: 992 bp Unfinished
sequence: d426N21 Contig_ID: 01102 Length: 38703 bp Unfinished
sequence: d426N21 Contig_ID: 02131 Length: 1244 bp.
***
*** WARNING: Phase 1 High Throughput Genome Sequence ***
***
* This sequence is unfinished. When sequencing is complete.
* the sequence data presented in this record will be replaced
* by a single finished sequence with the same accession number.
Location/Qualifiers
1. 117261
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="426N21"
/chromosome="X"
BASE COUNT 28333 a 27090 c 25467 g 27567 t 8804 others
ORIGIN
Query Match 5.9% Score 36.6; DB 18; Length 117261:
Best Local Similarity 55.0%; Pred. No. 2.1:
Matches 72; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
OY 491 aaacggtgttttcaagacatgagtgatgtgacctgcggaagaccacccctccctac 550
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88558 AGAATTGGCTATTTCACAGAGCGGAGGTTGTCAGTGAAGCCCATCGAGATCTGTCCTACT 88499
551 gaaccacactgtgagtgtgtgtgcacagtgcagcagacagaccttaaaaaaaaaaaga 610
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 88498 GCACCTCCAGCCTGGGTGTCACAGAGCGAGATTTCATACTAAAAAAAAAAAAAAAAAAAA 88439

```

QY 611 aacaacacaaa 621
Db 88438 AAAAAAAAAA 88428

RESULT 5
LOCUS HUM8DC62 3339 bp DNA PRI 22-AUG-1994
DEFINITION Homo sapiens (subclone H8 4_b9 from P1 35 H5 C8) DNA sequence.
ACCESSION L35661
NID 9532026
KEYWORDS MX1 region.
SOURCE Homo sapiens (library: subclones in pSP72 from P1 clone 35 H5 C8 (H8)) DNA.
ORGANISM Homo sapiens
Eukaryote; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 3339)
AUTHORS Martin,C.H., Bondoc,M.M., Chiang,A., Cloutier,T., Davis,C.A., Ericsson,C.L., Jaklevic,M.A., Kim,R.J., Lee,M.T., Li,M., Mayeda,C.A., Steiert-El Kheir,A. and Palazzolo,M.J.
TITLE Sequencing of the MX1 region on human chromosome 21
JOURNAL Unpublished (1994)
COMMENT Sequence submitted by:
Human Genome Center and
Drosophila Genome Center
Lawrence Berkeley Laboratory
Berkeley, CA 94720
e-mail: seq@genome.lbl.gov
This subclone overlaps H8 6_e2 and H8 3_d9.
FEATURES
source
1..3339
/organism="Homo sapiens"
/db_xref="taxon:9606"
/fisa_lib="Subclones in pSP72 from P1 clone 35 H5 C8 (H8)"

BASE COUNT 749 a 969 c 875 g 746 t
ORIGIN

Query Match 5.8%; Score 36; DB 10; Length 3339;
Best Local Similarity 49.0%; Pred. No. 2.2;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 289 gtccagaagatgagctggtgctatgagagatgacagatggtgcccgcagagcgca 348
Db 392 GTACATGATTAAGGAGATGATAGGTAGGACCCACAGTGGAGGCTCTGGAGGCGC 333
QY 349 ccggttcaaggagactgggtttccagaagtgtaaagcactgtgcgagactgtgcgtg 408
Db 332 CCACTGACAGCCAGCTCCAGGAGTCCAGAGTCCCGCTAGTGTGCTGCGTGGAGCGA 273
QY 409 gaacgccttcaaggagcactgctacacaccagtgtgctgtctgagggaactgct 468
Db 272 AATCTGTCTCCAGGAGACCTGCTCTCGGCTGCCAGCTGCCAAGTCAAGTAAGCT 213
QY 469 gccagatttacgg 484
Db 212 TTCAGAAATCTCACTG 197

RESULT 6
LOCUS AC005612 60904 bp DNA PRI 04-SEP-1998
DEFINITION Homo sapiens chromosome 21, P1 clone LBL#8 (UBNL H8), complete sequence.
ACCESSION AC005612
NID 93540153
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 60904)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Kadner,K., Miguel,T., Miller,C., Pitluck,S., Pollard,M., Rojsek,I., Sudramanian,S. and Martin,C.H.
TITLE Sequencing of human chromosome 21
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 60904)
AUTHORS Ricke,D.O.
TITLE Large Scale Sequence Analysis and Annotation with the Sequence Comparison Analysis (SCAN) System
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 60904)
AUTHORS Kimmerly,W., Bondoc,M., Cheng,J., Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T., Pitluck,S., Pollard,M., Rojsek,I., Sudramanian,S. and Martin,C.H.
TITLE Direct Submission
JOURNAL Submitted (04-SEP-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A.
COMMENT Sequence submitted by:
DOE Joint Genome Institute.
FEATURES
source
1..60904
/organism="Homo sapiens"
/db_xref="taxon:9606"
/map="21"
/clone="P1 LBL#8"
/chromosome="21"
/note="UNL H8"
1..675
/rpt_family="MER25"
423..565
/rpt_family="L1"
815..4661
/rpt_family="L1"
4917..4940
/note="(ACT)8"
/rpt_type=tandem
/rpt_unit=ACT
complement(3525..5890)
/rpt_family="TTR1"
complement(6688..6819)
/rpt_family="Alu"
7441..7741
/rpt_family="Alu"
7724..7743
/note="(A)20"
/rpt_type=tandem
/rpt_unit=A
complement(8063..8123)
/rpt_family="MER5"
complement(8811..9778)
/rpt_family="TTR5"
complement(9970..10409)
/rpt_family="MER41"
complement(11266..11344)
/rpt_family="TTR12"
12169..12204
/note="(GT)18"
/rpt_type=tandem
/rpt_unit=GT
complement(13074..13520)
/rpt_family="TTR7"
complement(13974..14029)
/rpt_family="MER5"
15399..15473
/rpt_family="L1"
15697..15808
/note="GRAIL 2 excellent exon, frame 0"
complement(15858..16136)
/rpt_family="Alu"
16565..16698
/note="GRAIL 2 excellent exon, frame 1"

```

repeat_region 17631..17863
/rpt_family="Alu"
repeat_region complement(18991..19253)
/rpt_family="Alu"
repeat_region complement(19424..19715)
/rpt_family="Alu"
misc_feature 20241..20286
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 21081..21263
/rpt_family="MER5"
repeat_region 21861..22013
/rpt_family="MER20"
repeat_region complement(23391..23728)
/rpt_family="THE1"
misc_feature 23779..23994
/note="GRAIL 2 excellent exon, frame 0"
repeat_region complement(24012..24087)
/rpt_family="MLT1"
misc_feature 24957..25094
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 25319..25585
/rpt_family="Alu"
misc_feature 27635..27825
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 28141..28430
/rpt_family="Alu"
misc_feature 28834..28966
/note="GRAIL 2 excellent exon, frame 0"
misc_feature 29686..29929
/note="GRAIL 2 excellent exon, frame 2"
repeat_region 29977..30000
/note="(GT)12"
/rpt_type=tandem
/rpt_unit=GT
/rpt_unit=GT
31701..31844
/note="GRAIL 2 excellent exon, frame 2"
repeat_region complement(31932..32057)
/rpt_family="MER5"
32297..32320
/note="(T)24"
/rpt_type=tandem
/rpt_unit=T
repeat_region complement(32303..32508)
/rpt_family="Alu"
33424..33584
/note="GRAIL 2 excellent exon, frame 0"
misc_feature 33976..34117
/note="GRAIL 2 excellent exon, frame 0"
repeat_region complement(34362..34558)
/rpt_family="T199er1"
repeat_region complement(34577..34696)
/rpt_family="T199er1"
repeat_region complement(35034..35111)
/rpt_family="MLT1"
repeat_region complement(35713..36022)
/rpt_family="Alu"
36501..36803
/rpt_family="Alu"
repeat_region 36909..36950
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
37127..37266
/note="GRAIL 2 excellent exon, frame 2"
repeat_region 39543..39832
/rpt_family="Alu"
40592..40840
/note="GRAIL 2 excellent exon, frame 1"
repeat_region complement(40971..41337)
/rpt_family="THE1"
repeat_region complement(41617..41717)
/rpt_family="MER4"
repeat_region complement(42881..43215)

```

```

repeat_region /rpt_family="Alu"
43942..44219
/rpt_family="Alu"
repeat_region complement(44724..45031)
/rpt_family="Alu"
repeat_region complement(46007..46178)
/rpt_family="Alu"
misc_feature 46500..46725
/note="GRAIL 2 excellent exon, frame 2"
repeat_region 47500..47781
/rpt_family="Alu"
misc_feature complement(49242..49309)
/note="GRAIL 2 excellent exon, frame 2"
repeat_region 49335..50077
/rpt_family="THE1"
repeat_region 50080..51967
/rpt_family="MSTAR"
repeat_region complement(50244..50543)
/rpt_family="Alu"
misc_feature complement(51595..51645)
/note="GRAIL 2 excellent exon, frame 1"
repeat_region 51968..52329
/rpt_family="THE1"
repeat_region 55080..55378
/rpt_family="Alu"
misc_feature complement(56369..56511)
/note="GRAIL 2 excellent exon, frame 0"
misc_feature complement(58621..58678)
/note="GRAIL 2 excellent exon, frame 0"
repeat_region 59339..59470
/rpt_family="MER20"
misc_feature complement(59793..59910)
/note="GRAIL 2 excellent exon, frame 1"
BASE COUNT 16756 a 14066 c 14518 g 15564 t
ORIGIN

```

Query Match 5.8%; Score 36; DB 12; Length 60904;
 Best Local Similarity 49.0%; Pred. No. 3;
 Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

```

QY 289 gtccaagaaatgtgcttgcgcctatggagagatgcacagtgtgtccctgcagccgca 348
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16307 GTCAATGATTTAAAGGAGATGATGAGGTAAGGACCCACAGTGGAGGCGCAGC 16366

QY 349 ccggttcaaggaagactgtggtttccagaagttaagcaatgtgcggaactgtgcctgt 408
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16367 CCACTACACGCCACTCCAGGAGCTCCAGAGTCCCTCTAGTGTGGTGTGAGGGA 16426

QY 409 gaaccgcttccagagggccaactgtctcacaccagtgatgtctgtcgcgggaactgtcct 468
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16427 AATCTTTCCTCCAGGAGACTGTCTCGGCTCCAGCTGCCAAGTCAAGATAAGCT 16486

QY 469 gccaggaatttaccgg 484
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 16487 TTCAGAAATCTCACTG 16502

```

```

RESULT 7
HSAJ11929_5
WPCOMMENT
Sequence split into 7 fragments LOCUS HSAJ11929 Accession AJ011929
Fragment Name Begin End
HSAJ11929_0 1 210000
HSAJ11929_1 100001 210000
HSAJ11929_2 200001 310000
HSAJ11929_3 300001 410000
HSAJ11929_4 400001 510000
HSAJ11929_5 500001 610000
HSAJ11929_6 600001 614109
Continuation (6 of 7) of HSAJ11929 from base 500001 (AJ011929 Homo sapiens *** SEQUEN

```

[illegible]

| RESULT | 9 | | | | |
|------------|--|--|--|--|--|
| LOCUS | AC005612 | | | | |
| DEFINITION | Homo sapiens chromosome 21, p1 clone LBL#8 (LBL H8), complete sequence. | | | | |
| ACCESSION | AC005612 | | | | |
| NID | 93540153 | | | | |
| KEYWORDS | HTG. | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. | | | | |
| AUTHORS | 1 (bases 1 to 60904) Kimmerly,W., Bondoc,M., Cheng,D., Connolly,K.S., Gunning,K.M., Kadner,K., Miguel,T., Miller,C., Pfluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H. | | | | |
| TITLE | Sequencing of human chromosome 21 | | | | |
| JOURNAL | Unpublished | | | | |
| REFERENCE | 2 (bases 1 to 60904) Ricke,D.O. | | | | |
| AUTHORS | Large Scale Sequence Analysis and Annotation with the Sequence Comparison Analysis (SCAN) System | | | | |
| TITLE | Unpublished | | | | |
| JOURNAL | 3 (bases 1 to 60904) | | | | |
| REFERENCE | Kimmerly,W., Bondoc,M., Cheng,D., Connolly,K.S., Gunning,K.M., Davis,C.A., Kadner,K., Miguel,T., Pfluck,S., Pollard,M., Rojeski,H., Subramanian,S. and Martin,C.H. | | | | |
| AUTHORS | Direct Submission | | | | |
| TITLE | Submitted (04-SEP-1998) Human Genome Center, DOE Joint Genome Institute, Lawrence Berkeley National Laboratory, MS 74-157, Berkeley, CA 94720, U.S.A. | | | | |
| JOURNAL | Sequence submitted by: | | | | |
| REFERENCE | DOE Joint Genome Institute. | | | | |
| COMMENT | Location/Qualifiers | | | | |
| FEATURES | 1. 60904 | | | | |
| SOURCE | /organism="Homo sapiens" /db_xref="taxon:9606" /map="21" /clone="P1 LBL#8" /chromosome="21" /note="LBL H8" 1. 675 /rpt_family="MER25" 423. 565 /rpt_family="L1" 815. 4661 /rpt_family="L1" 4917. 4940 /note="(ACT)8" /rpt_type=tandem /rpt_unit=ACT complement(5525..5890) /rpt_family="THER1" complement(6688..6819) /rpt_family="Alu" 7441..7741 /rpt_family="Alu" 7724..7743 /note="(A)20" /rpt_type=tandem /rpt_unit=A complement(8063..8123) /rpt_family="MER5" complement(8811..9778) /rpt_family="LTR5" complement(9970..10409) /rpt_family="MER41" complement(11266..11344) /rpt_family="LTR12" | | | | |


```
repeat_region 12169..12204
/note="(GT)18"
/rpt_type=tandem
/rpt_unit=GT
repeat_region complement(13074..13520)
/rpt_family="MER7"
repeat_region complement(13974..14029)
/rpt_family="MERS"
repeat_region 15399..15473
/rpt_family="L1"
misc_feature 15697..15808
/note="GRAIL 2 excellent exon, frame 0"
complement(15858..16136)
/rpt_family="Alu"
misc_feature 16565..16698
/note="GRAIL 2 excellent exon, frame 1"
17631..17863
/rpt_family="Alu"
repeat_region complement(18991..19253)
/rpt_family="Alu"
repeat_region complement(19424..19715)
/rpt_family="Alu"
20241..20286
/note="GRAIL 2 excellent exon, frame 0"
21081..21263
/rpt_family="MERS"
21861..22013
/rpt_family="MER20"
repeat_region complement(23391..23728)
/rpt_family="THE1"
23779..23994
/note="GRAIL 2 excellent exon, frame 0"
complement(24012..24087)
/rpt_family="MLT1"
24957..25094
/note="GRAIL 2 excellent exon, frame 1"
25319..25585
/rpt_family="Alu"
27635..27825
/note="GRAIL 2 excellent exon, frame 0"
28141..28430
/rpt_family="Alu"
28834..28966
/note="GRAIL 2 excellent exon, frame 0"
29686..29929
/note="GRAIL 2 excellent exon, frame 2"
29977..30000
/note="(GT)12"
/rpt_type=tandem
/rpt_unit=GT
31701..31844
/note="GRAIL 2 excellent exon, frame 2"
complement(31932..32057)
/rpt_family="MERS"
32297..32320
/note="(T)24"
/rpt_type=tandem
/rpt_unit=T
complement(32303..32508)
/rpt_family="Alu"
33424..33584
/note="GRAIL 2 excellent exon, frame 0"
33976..34117
/note="GRAIL 2 excellent exon, frame 0"
complement(34362..34558)
/rpt_family="trigger1"
complement(34577..34696)
/rpt_family="trigger1"
complement(35034..35111)
/rpt_family="MLT1"
complement(35713..36022)
/rpt_family="Alu"
36501..36803
```

```
repeat_region /rpt_family="Alu"
36909..36950
/note="(GT)21"
/rpt_type=tandem
/rpt_unit=GT
misc_feature 37127..37266
/note="GRAIL 2 excellent exon, frame 2"
39543..39822
/rpt_family="Alu"
misc_feature 40592..40840
/note="GRAIL 2 excellent exon, frame 1"
complement(40971..41337)
/rpt_family="THE1"
repeat_region complement(41617..41717)
/rpt_family="MER4"
repeat_region complement(42881..43215)
/rpt_family="Alu"
43942..44219
/rpt_family="Alu"
repeat_region complement(44724..45031)
/rpt_family="Alu"
repeat_region complement(46007..46178)
/rpt_family="Alu"
46500..46725
/note="GRAIL 2 excellent exon, frame 2"
47500..47781
/rpt_family="Alu"
misc_feature complement(49242..49309)
/note="GRAIL 2 excellent exon, frame 2"
49935..50077
/rpt_family="THE1"
50080..51967
/rpt_family="MSTAR"
repeat_region complement(50244..50543)
/rpt_family="Alu"
misc_feature complement(51595..51645)
/note="GRAIL 2 excellent exon, frame 1"
51968..52329
/rpt_family="THE1"
repeat_region 55080..55378
/rpt_family="Alu"
misc_feature complement(56369..56511)
/note="GRAIL 2 excellent exon, frame 0"
complement(58621..58678)
/note="GRAIL 2 excellent exon, frame 0"
59339..59470
/rpt_family="MER20"
misc_feature complement(59793..59910)
/note="GRAIL 2 excellent exon, frame 1"
BASE COUNT 16756 a 14066 c 14518 g 15564 t
ORIGIN

Query Match 5.88; Score 36; DB 42; Length 60904;
Best Local Similarity 49.08; Pred. No. 3;
Matches 96; Conservative 0; Mismatches 100; Indels 0; Gaps 0;

QY 289 gtccaaggaatgctgctgctatgagggaagatgacagatgctgctgacgacga 348
Db 16307 GTACATGATTAAAGGAGGTATAGGTAAAGACCAAGTGAAGCTCTGGAGCCAGC 16366
QY 349 ccggttccaaggaagactggtttccaaagttaagcgaatgctgacgtgtg 408
Db 16367 CCACGTGACAGCCACTCCAGGAGAGTCCAAATGCCCTAGTGTGGTGTGAGGGA 16426
QY 409 gaacgccttcaaggagcgaactctcacacacgaatgctgctgcygggacct 468
Db 16427 AATCTGTTCTCCACAGGAGCTCTGCTGCTGCCACAGCTGCAAAAGTCAAGAAATTAAGCT 16486
QY 469 gccaggaatttaccgg 484
Db 16487 TTCAAAATCTCACTG 16502
```

| | |
|-----------------------|---|
| NID | g2673898 |
| KEYWORDS | Htg. |
| SOURCE | human. |
| ORGANISM | Homo sapiens Eukaryote; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo. 1 (bases 1 to 140977) |
| REFERENCE | Evans,G.A., Athanasiou,M., Basil,M., Bradbury,P., Brignac,S., Bunester,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hanner,L., Harris,J., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R. template |
| TITLE | Unpublished |
| JOURNAL | 2 (bases 1 to 140977) |
| REFERENCE | Evans,G.A., Bradbury,P., Brignac,S., Bunester,R., Burbee,D., Davis,C., Davies,C.J., Davis,C., English,C., Fondon,T., Franklin,T.L., Garner,H.R., Gordon,M., Gotway,G., Grant,O., Hanner,L., Harris,J., Hinson,S., Megarity,C., Narayanaswamy,U., Newton,J., O'Brien,K., Oliver,T., Patel,P., Probst,S., Rayner,S., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R. Direct Submission |
| TITLE | Submitted (11-SEP-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA 3 (bases 1 to 140977) |
| REFERENCE | Evans,G.A., Athanasiou,M., Basil,M., Bradbury,P., Brignac,S., Gee,V., Gordon,M., Gotway,G., Grant,O., Hanner,L., Harris,J., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R. Direct Submission |
| TITLE | Submitted (10-DEC-1997) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA On Dec 10, 1997 this sequence version replaced g1:2392758. |
| JOURNAL | |
| COMMENT | |
| FEATURES | |
| source | Location/Qualifiers 1..140977 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="PD1075f20" /chromosome="11" |
| BASE COUNT | 34355 a 37819 c 37235 g 31568 t |
| ORIGIN | |
| Query Match | 5.7%, Score 35.6, DB 11; Length 140977; |
| Best Local Similarity | 50.0%; Pred. No. 4.3; |
| Matches | 89; Conservative 0; Mismatches 89; Indels 0; Gaps 0. |
| OY | 220 gcagcagaatccaagtgcgatctcggaactgttcccttcycaaacagtcgcgacctyg 279 |
| Db | 55124 GGAGGAGCAATTTCCACGACCAGGGCGCAGACGAGCACGCCCTGTTCAGCATCTGGAGAGG 55183 |
| OY | 280 catgagctgtccaagagatgtgcctcgcatgaggagagatgcagagctgtgtccctg 339 |
| Db | 55184 CCGGAGTGTCTCCGAATGATGATTACCTGGAAACATTTTCACCCCTTTAGGCGCCAC 55243 |
| OY | 340 caggcgacaccggttccaagagaaactgtgggtttccaagaagtgtaaagcatgtgagac 397 |
| Db | 55244 CCTGCACAGAGGCCCGACGAGAACCCCTCTATGTGGGGCGCGACGCGCTGCCCCAC 55301 |
| RESULT | 12 |
| LOCUS | AC002536 140977 bp DNA PRI 10-DEC-1997 |
| DEFINITION | Human Chromosome 11 pac pD1075f20, complete sequence. |
| ACCESSION | AC002536 |
| NID | g2673898 |
| KEYWORDS | Htg. |


```

/rpt_family="Alusq"
repeat_region
6988. .7007
/rpt_family="(TAAA)n"
repeat_region
7141. .7173
/rpt_family="AT_rich"
repeat_region
complement(7360. .7633)
/rpt_family="Tiggers3(Golem)"
repeat_region
7634. .7942
/rpt_family="Alusx"
complement(7643. .8051)
/rpt_family="Tiggers3(Golem)"
repeat_region
complement(8528. .8838)
/rpt_family="Alusx"
repeat_region
complement(8901. .9240)
/rpt_family="Alub"
repeat_region
10142. .10422
/rpt_family="Alusq"
repeat_region
10473. .10749
/rpt_family="Alub"
repeat_region
complement(10750. .10769)
/rpt_family="(CAAAA)n"
repeat_region
complement(10771. .11044)
/rpt_family="Alub"
repeat_region
11249. .11366
/rpt_family="Alusq/X"
repeat_region
11371. .11391
/rpt_family="AT_rich"
repeat_region
11417. .11703
/rpt_family="Alub"
repeat_region
11888. .12188
/rpt_family="Alusq"
repeat_region
12189. .12263
/rpt_family="MER41B"
repeat_region
12275. .12290
/rpt_family="Alub"
repeat_region
12291. .12582
/rpt_family="Alusq"
repeat_region
12583. .12731
/rpt_family="Alub"
repeat_region
13548. .13659
/rpt_family="Alub"
repeat_region
15721. .16004
/rpt_family="Alub"
repeat_region
16307. .17274
/rpt_family="Alub"
repeat_region
17651. .17753
/rpt_family="LIMS"
repeat_region
complement(17754. .18051)
/rpt_family="Alusq"
repeat_region
18052. .18063
/rpt_family="LIMS/D"
repeat_region
18167. .18481
/rpt_family="Alusx"
repeat_region
18486. .18796
/rpt_family="Alub"
repeat_region
18846. .18873
/rpt_family="AT_rich"
repeat_region
18945. .19239
/rpt_family="Alusx"
repeat_region
19243. .19509
/rpt_family="Alub"
repeat_region
complement(19558. .19700)
/rpt_family="L2"
repeat_region
19709. .19968
/rpt_family="HAL1"
repeat_region
20013. .20318
/rpt_family="Alusq"
repeat_region
20343. .20369
/rpt_family="AT_rich"
complement(20370. .20495)
/rpt_family="FIAM.C"
complement(20502. .20797)
/rpt_family="Alusq"

repeat_region
20798. .20815
/rpt_family="AT_rich"
repeat_region
20866. .21093
/rpt_family="HAL1"
repeat_region
21094. .21398
/rpt_family="Alub"
repeat_region
21405. .21715
/rpt_family="Alusx"
repeat_region
21716. .21785
/rpt_family="HAL1"
repeat_region
21791. .22093
/rpt_family="Alusq"
repeat_region
22187. .22465
/rpt_family="Alusq1"
repeat_region
22466. .22521
/rpt_family="(CAAAA)n"
repeat_region
22643. .22938
/rpt_family="Alusx"
repeat_region
22947. .23251
/rpt_family="Alusq"
repeat_region
23784. .24097
/rpt_family="Alub"
repeat_region
24698. .24993
/rpt_family="Alusq"
repeat_region
25122. .25147
/rpt_family="(CA)n"
repeat_region
25698. .26013
/rpt_family="Alusq"
repeat_region
26078. .26104
/rpt_family="AT_rich"
repeat_region
complement(27061. .27091)
/rpt_family="(CA)n"
repeat_region
28173. .28481
/rpt_family="Alusq"
repeat_region
28829. .29117
/rpt_family="Alusq"
repeat_region
29407. .29727
/rpt_family="Alusq"
repeat_region
complement(2974. .30261)
/rpt_family="Alusq"
repeat_region
31297. .31320
/rpt_family="AT_rich"
repeat_region
complement(31353. .31636)
/rpt_family="Alusq"
repeat_region
31885. .31907
/rpt_family="AT_rich"
repeat_region
32571. .32718
/rpt_family="Alusx"
repeat_region
32824. .32994
/rpt_family="Alusx"
repeat_region
33705. .33731
/rpt_family="GC_rich"
repeat_region
complement(33856. .33971)
/rpt_family="(CCGCG)n"
repeat_region
34016. .34043
/rpt_family="GC_rich"
repeat_region
complement(34842. .35072)
/rpt_family="Alusx"
repeat_region
35111. .35419
/rpt_family="Alusq"
repeat_region
35420. .35517
/rpt_family="purine-rich"
repeat_region
35518. .35638
/rpt_family="(GAA)n"
complement(36572. .36603)
/rpt_family="L2"
repeat_region
36626. .36741
/rpt_family="FIAM.C"
repeat_region
36779. .36821
/rpt_family="MER7"
complement(36849. .36877)
/rpt_family="(TAAA)n"
repeat_region
36892. .37008

```

```
repeat_region      /rpt_family="AluSg/x"
complement(37014..37315)
repeat_region      /rpt_family="AluSp"
complement(37316..37618)
repeat_region      /rpt_family="AluSx"
complement(37643..37671)
repeat_region      /rpt_family="(CAA)n"
complement(37672..37955)
repeat_region      /rpt_family="AluUb"
complement(38073..38228)
repeat_region      /rpt_family="AluSp/q"
38377..38500
repeat_region      /rpt_family="FLAM_C"
38501..38528
repeat_region      /rpt_family="(GAA)n"
38529..38630
repeat_region      /rpt_family="FLAM_C"
38639..38926
repeat_region      /rpt_family="AluSx"
39001..39036
repeat_region      /rpt_family="AT-rich"
complement(39115..39152)
repeat_region      /rpt_family="(GAA)n"
39176..39471
repeat_region      /rpt_family="AluSx"
39486..39644
repeat_region      /rpt_family="MLTIF"
complement(39673..39963)
repeat_region      /rpt_family="L1PB1"
39970..40272
repeat_region      /rpt_family="AluSp"
41835..41868
repeat_region      /rpt_family="POLY_A"
41871..42183
repeat_region      /rpt_family="AluY"
complement(42184..42379)
repeat_region      /rpt_family="L1PB1"
42380..42671
repeat_region      /rpt_family="AluSc"
complement(42672..42691)
repeat_region      /rpt_family="L1PB1"
42754..43049
repeat_region      /rpt_family="MLTIF"
complement(43054..43356)
repeat_region      /rpt_family="AluUo"
43363..43393
repeat_region      /rpt_family="AT-rich"
complement(43400..43414)
repeat_region      /rpt_family="MTR"
complement(43415..43723)
repeat_region      /rpt_family="AluSx"
complement(43724..43884)
repeat_region      /rpt_family="MTR"
complement(44278..44464)
repeat_region      /rpt_family="MLT1A1"
complement(44465..44774)
repeat_region      /rpt_family="AluSx"
complement(44775..44909)
repeat_region      /rpt_family="MLT1A1"
44914..45213
repeat_region      /rpt_family="AluY"
45223..45527
repeat_region      /rpt_family="AluSp"
45620..45752
repeat_region      /rpt_family="AluUb"
45753..46028
repeat_region      /rpt_family="AluSx"
46031..46055
repeat_region      /rpt_family="(CAAA)n"
46056..46342
repeat_region      /rpt_family="AluSg"
46343..46421
repeat_region      /rpt_family="AluUb"

repeat_region      46451..46500
/rpt_family="(TAA)n"
complement(46504..46697)
repeat_region      /rpt_family="HERVL"
complement(46712..46811)
repeat_region      /rpt_family="MERVL"
complement(47100..48514)
repeat_region      /rpt_family="SVA"
48524..48673
repeat_region      /rpt_family="AluSg/x"
complement(49245..49330)
repeat_region      /rpt_family="MER20"
49331..49629
repeat_region      /rpt_family="AluSp"
complement(49630..49711)
repeat_region      /rpt_family="MER20"
complement(49773..49870)
repeat_region      /rpt_family="AluUo/FRAM"
49952..50063
repeat_region      /rpt_family="purine-rich"
complement(50064..50135)
repeat_region      /rpt_family="AluU/FLAM"
complement(50136..50563)
repeat_region      /rpt_family="MLT2A"
50658..50789
repeat_region      /rpt_family="AluSg/x"
50792..51074
repeat_region      /rpt_family="AluSg"
complement(51313..51533)
repeat_region      /rpt_family="L2"
complement(51532..51673)
repeat_region      /rpt_family="AluSg/x"
51674..51713
repeat_region      /rpt_family="LIME1"
51714..52008
repeat_region      /rpt_family="MER2"
52009..52068
repeat_region      /rpt_family="LIME1"
52069..52361
repeat_region      /rpt_family="AluSx"
52362..52652
repeat_region      /rpt_family="LIME1"
complement(53828..53857)
repeat_region      /rpt_family="(CAA)n"
complement(53858..54129)
repeat_region      /rpt_family="AluSx"
54336..54464
repeat_region      /rpt_family="L2"
complement(54679..54876)
repeat_region      /rpt_family="MTR"
complement(54955..55100)
repeat_region      /rpt_family="polypurine"
57136..57172
repeat_region      /rpt_family="GC-rich"
57259..57279
repeat_region      /rpt_family="GC-rich"
complement(57293..57683)
repeat_region      /rpt_family="purine-rich"
complement(57790..57927)
repeat_region      /rpt_family="L2"
complement(58042..58151)
repeat_region      /rpt_family="MER21B"
58233..58473
repeat_region      /rpt_family="AluUb"
58474..58752
repeat_region      /rpt_family="AluSg1"
complement(58929..59240)
repeat_region      /rpt_family="AluUo"
complement(59340..59645)
repeat_region      /rpt_family="AluSp"
complement(59652..59791)
repeat_region      /rpt_family="FLAM_C"
complement(59794..59841)
```

```

repeat_region /rpt_family="(TAAA)n"
complement(59842..60119)
repeat_region /rpt_family="Alusx"
60120..60145
/rpt_family="AT_rich"
repeat_region complement(60269..60398)
/rpt_family="MLT1A2"
repeat_region complement(60399..60502)
/rpt_family="MER6B"
complement(60503..60788)
/rpt_family="Alusx"
repeat_region complement(60789..61024)
/rpt_family="MER6B"
61025..61312
/rpt_family="Alusg"
complement(61313..61485)
/rpt_family="MER6B"
repeat_region complement(61486..61639)
/rpt_family="MLT1A2"
complement(61647..61815)
/rpt_family="Alusx"
repeat_region complement(61816..62123)
/rpt_family="AluYb"
complement(62124..62275)
/rpt_family="Alusx"
62276..62545
/rpt_family="L1M4"
repeat_region 62571..62662
/rpt_family="Alusg"
complement(62663..63313)
/rpt_family="SVA"
63314..63369
/rpt_family="Alusg"
63370..63601
/rpt_family="Alusg"
63604..63708
/rpt_family="L1M4"
complement(63785..64085)
/rpt_family="AluY"
64101..64161
/rpt_family="L1M4"
complement(64162..64333)
/rpt_family="FAM"
64334..64837
/rpt_family="L1M4"
complement(64838..64988)
/rpt_family="AluYb"
complement(64989..65282)
/rpt_family="AluSp"
complement(65283..65430)
/rpt_family="AluYb"
65431..65646
/rpt_family="L1M4"
complement(65647..65815)
/rpt_family="AlusC"
complement(65816..66116)
/rpt_family="AluSp"
complement(66117..66161)
/rpt_family="AlusC"
66162..66229
/rpt_family="L1M4"
66232..66302
/rpt_family="L1ME"
66325..66368
/rpt_family="Alusg/x"
66369..66671
/rpt_family="Alusx"
66672..66852
/rpt_family="Alusg/x"
66909..66915
/note="Single-stranded coverage."
66912..67000
/rpt_family="HAL1"

repeat_region 67001..67305
/rpt_family="Alusx"
67306..67363
/rpt_family="HAL1"
repeat_region complement(67372..67530)
/rpt_family="FAM"
67531..67688
/rpt_family="AluYb/FLAM"
complement(67689..68034)
/rpt_family="Alusx"
repeat_region complement(68036..68328)
/rpt_family="Alusg"
complement(68427..68721)
/rpt_family="Alusg"
68750..69186
/rpt_family="L2"
69247..69346
/rpt_family="MER5A"
complement(69347..69551)
/rpt_family="Alusg"
69552..69859
/rpt_family="AluY"
complement(69860..69985)
/rpt_family="Alusg"
69986..70026
/rpt_family="MER5A"
70172..70287
/rpt_family="L2"
complement(70291..70564)
/rpt_family="AluYb"
70566..70626
/rpt_family="Alu"
70690..70771
/rpt_family="HAL1"
70779..70807
/rpt_family="Alu"
71156..71189
/rpt_family="(TA)n"
complement(71249..71356)
/rpt_family="MIR"
complement(71550..71586)
/rpt_family="tRNA-Tyr-TAC"
complement(71616..71906)
/rpt_family="AluYb"
72074..72106
/rpt_family="L2"
72174..72205
/rpt_family="MER4D"
72212..72233
/rpt_family="AT_rich"
complement(72310..72609)
/rpt_family="Alusg"
complement(72863..72880)
/rpt_family="AluYb"
complement(72881..72917)
/rpt_family="(CA)n"
complement(72918..73202)
/rpt_family="AluYb"
73232..73537
/rpt_family="AluY"
complement(73542..73635)
/rpt_family="MIR"
complement(73717..73994)
/rpt_family="Alusg"
complement(74083..74393)
/rpt_family="Alusx"
74529..74717
/rpt_family="L2"
75055..75100
/rpt_family="AT_rich"
complement(75524..75896)
/rpt_family="L1M3e"
76331..76355

```

repeat_region /rpt_family="AT_rich"
78004. .78036
repeat_region /rpt_family="AT_rich"
79901. .80186
repeat_region /rpt_family="Alusx"
80340. .80372
repeat_region /rpt_family="AT_rich"
80373. .80668
repeat_region /rpt_family="AluYb8"
complement(80759. .81066)
repeat_region /rpt_family="AluY"
81223. .81533
repeat_region /rpt_family="Alusg"
complement(81615. .81909)
repeat_region /rpt_family="AlusC"
81911. .82058
repeat_region /rpt_family="MIR"
complement(82092. .82382)
repeat_region /rpt_family="Alusx"
82445. .82612
repeat_region /rpt_family="MIR"
complement(82641. .82805)
repeat_region /rpt_family="MLT1B"
complement(82848. .83145)
repeat_region /rpt_family="Alusx"
complement(83146. .83281)
repeat_region /rpt_family="AluYb"
complement(83282. .83404)
repeat_region /rpt_family="MLT1B"
complement(83405. .83694)
repeat_region /rpt_family="Alusx"
complement(83695. .83756)
repeat_region /rpt_family="MLT1B"
complement(83757. .84065)
repeat_region /rpt_family="AluJo"
complement(84066. .84114)
repeat_region /rpt_family="MLT1B"
complement(84272. .84557)
repeat_region /rpt_family="Alusg"
84563. .84818
repeat_region /rpt_family="L2"
complement(84819. .85129)
repeat_region /rpt_family="Alusx"
85130. .85480
repeat_region /rpt_family="L2"
85717. .85816
repeat_region /rpt_family="HY3"
complement(86213. .86516)
repeat_region /rpt_family="AlusC"
complement(86959. .87253)
repeat_region /rpt_family="Alusx"
87281. .87573
repeat_region /rpt_family="AluYb8"
87574. .87599
repeat_region /rpt_family="(CAA)n"
87600. .87612
repeat_region /rpt_family="AluYb8"
87891. .88354
repeat_region /rpt_family="L1MC/D"
88370. .88663
repeat_region /rpt_family="Alusg"
complement(89376. .89597)
repeat_region /rpt_family="MER30"
complement(90386. .90696)
repeat_region /rpt_family="Alusg"
91095. .91396
repeat_region /rpt_family="Alusx"
complement(91689. .91997)
repeat_region /rpt_family="Alusg"
complement(92020. .92131)
repeat_region /rpt_family="Alusg/x"
complement(92191. .92249)
repeat_region /rpt_family="MLT1A2"

repeat_region complement(92251. .92558)
repeat_region /rpt_family="AluJo"
complement(92745. .93059)
repeat_region /rpt_family="Alusg"
complement(93060. .93192)
repeat_region /rpt_family="AluYb"
complement(93193. .93293)
repeat_region /rpt_family="L1M4"
complement(93294. .93353)
repeat_region /rpt_family="AluJo"
complement(93354. .93658)
repeat_region /rpt_family="Alusg1"
complement(93659. .93867)
repeat_region /rpt_family="AluJo"
complement(93868. .94306)
repeat_region /rpt_family="L1M4"
complement(94312. .94443)
repeat_region /rpt_family="FLAM_A"
complement(94445. .94777)
repeat_region /rpt_family="L1M4"
complement(94778. .95092)
repeat_region /rpt_family="Alusg"
complement(95093. .95189)
repeat_region /rpt_family="L1M4"
complement(95190. .95488)
repeat_region /rpt_family="Alusx"
complement(95489. .96317)
repeat_region /rpt_family="L1M4"
complement(96413. .96713)
repeat_region /rpt_family="Alusx"
complement(96714. .96992)
repeat_region /rpt_family="AluJo"
complement(97632. .97938)
repeat_region /rpt_family="Alusg"
complement(98515. .98684)
repeat_region /rpt_family="MER46B"
complement(98685. .98866)
repeat_region /rpt_family="AluJo"
complement(98867. .99159)
repeat_region /rpt_family="AluYb"
complement(99160. .99298)
repeat_region /rpt_family="AluJo"
complement(99299. .99480)
repeat_region /rpt_family="MER46B"
99677. .99978
repeat_region /rpt_family="Alusx"
100115. .100142
repeat_region /rpt_family="AT_rich"
complement(100143. .100427)
repeat_region /rpt_family="AluY"
100428. .100445
repeat_region /rpt_family="AT_rich"
complement(100822. .100849)
repeat_region /rpt_family="(CA)n"
complement(102147. .102436)
repeat_region /rpt_family="Alusx"
complement(102439. .102465)
repeat_region /rpt_family="POLY_A"
complement(102608. .102751)
repeat_region /rpt_family="MIR"
complement(102756. .102859)
repeat_region /rpt_family="L2"
complement(103236. .103527)
repeat_region /rpt_family="Alusx"
complement(103725. .103756)
repeat_region /rpt_family="polypurine"
complement(103757. .104063)
repeat_region /rpt_family="AluJo"
complement(104625. .104928)
repeat_region /rpt_family="AluYb"
105057. .105317
repeat_region /rpt_family="AluJo"
106712. .106734

```
repeat_region /rpt_family="GC_rich"
107291. .107362
/rpt_family="GC_rich"
complement(107433. .107457)
repeat_region /rpt_family="GCAG)n"
107595. .107618
/rpt_family="GC_rich"
107733. .107762
/rpt_family="GC_rich"
107928. .108001
/rpt_family="GC_rich"
complement(108073. .108129)
/rpt_family="purine-rich"
complement(110962. .111287)
/rpt_family="AluSx"
complement(111508. .111578)
/rpt_family="11M3e"
complement(111639. .111804)
/rpt_family="FAM"
complement(111827. .112139)
/rpt_family="AluSx"
112907. .112928
/rpt_family="AT_rich"
113857. .114160
/rpt_family="AluSx"
complement(114487. .114794)
/rpt_family="AluYb8"
complement(115026. .115147)
/rpt_family="FRAM_C"
complement(115154. .115233)
/rpt_family="(CA)n"
complement(115676. .115761)
/rpt_family="L2"
115762. .116050
/rpt_family="AluSg"
complement(116051. .116264)
/rpt_family="L2"
11638. .116700
/rpt_family="L2"
complement(116960. .117219)
/rpt_family="AluSx"
complement(117220. .117355)
/rpt_family="G-rich"
complement(117336. .117590)
/rpt_family="AluYb"
complement(117632. .117741)
/rpt_family="FRAM_A"
117742. .117809
/rpt_family="L2"
118052. .118266
/rpt_family="MIR"
complement(118930. .119080)
/rpt_family="L2"
119157. .119454
/rpt_family="AluSg"
complement(119486. .119609)
/rpt_family="L2"
119798. .119961
/rpt_family="L2"
119993. .120040
/rpt_family="(TA)n"
complement(120042. .120170)
/rpt_family="FRAM_C"
complement(120180. .120356)
/rpt_family="AluSg/x"
complement(120360. .120674)
/rpt_family="AluSg"
complement(120809. .121101)
/rpt_family="AluSg"
121680. .121893
/rpt_family="MER20"
122812. .123110
/rpt_family="AluYb8"

repeat_region complement(123636. .123802)
repeat_region /rpt_family="MIR"
complement(124180. .124479)
/rpt_family="AluSg"
125035. .125070
/rpt_family="AT_rich"
125239. .125524
/rpt_family="AluSx"
125525. .125577
/rpt_family="(CA)n"
complement(125807. .125875)
/rpt_family="Charlie1"
125903. .125924
/rpt_family="AT_rich"
125933. .126199
/rpt_family="AluYb"
complement(126217. .126407)
/rpt_family="Charlie1a"
126408. .126681
/rpt_family="AluSc"
complement(126682. .127379)
/rpt_family="Charlie1a"
complement(127634. .128101)
/rpt_family="L2"
complement(130705. .130741)
/rpt_family="(CAAA)n"
complement(130746. .130989)
/rpt_family="AluSg"
132238. .132536
/rpt_family="AluSp"
complement(133820. .133882)
/rpt_family="A-rich"
complement(134361. .134658)
/rpt_family="AluSg"
complement(134660. .134959)
/rpt_family="AluSx"
135078. .135375
/rpt_family="AluSx"
135378. .135669
/rpt_family="AluY"
complement(135741. .135752)
/rpt_family="AluSg/x"
complement(135753. .135779)
/rpt_family="(CAT)n"
complement(135780. .135944)
/rpt_family="AluSg/x"
136851. .136881
/rpt_family="MIR"
137363. .137522
/rpt_family="FRAM"
complement(137618. .138412)
/rpt_family="L2"
138422. .138444
/rpt_family="AT_rich"
complement(139440. .139755)
/rpt_family="AluY"
complement(139820. .140012)
/rpt_family="MIR"
140299. .140579
/rpt_family="AluYb"
complement(140756. .140818)
/rpt_family="MIR"
142048. .142358
/rpt_family="AluSg"
142359. .142524
/rpt_family="AluSg/x"
complement(143972. .144135)
/rpt_family="AluSg/x"
144187. .144263
/rpt_family="FRAM"
144276. .144298
/rpt_family="AT_rich"
complement(144535. .144820)
```



```

repeat_region      /rpt_family="AluSx"      complement(144945..145101)
                    /rpt_family="FRAM"
repeat_region      /rpt_family="MER20" complement(145210..145403)
                    /rpt_family="AluSp"
repeat_region      /rpt_family="FLAM_C" 145934..146027
                    /rpt_family="AluCb" 146309..146538
                    /rpt_family="AluDb" 148821..148988
                    /rpt_family="MER5B" complement(149198..149678)
                    /rpt_family="L2" complement(150807..150838)
                    /rpt_family="AluYb" complement(150839..150864)
                    /rpt_family="TAAAn" complement(150865..151162)
                    /rpt_family="AluYb" 151247..151464
                    /rpt_family="MER58A" 151588..151891
                    /rpt_family="AluY" 151970..151990
                    /rpt_family="AT-rich" 152003..152297
                    /rpt_family="AluSx" complement(152753..152937)
                    /rpt_family="MR" complement(153254..153291)
                    /rpt_family="(GA)n" complement(153292..153595)
                    /rpt_family="AluYb" 153812..153834
                    /rpt_family="AT-rich" complement(153843..153965)
                    /rpt_family="FLAM_A" complement(154028..154336)
                    /rpt_family="AluYb" 154373..154415
                    /rpt_family="MER66B" 154461..154624
                    /rpt_family="MR" 154864..155159
                    /rpt_family="AluSx" 156920..157056
                    /rpt_family="MR" complement(157057..157365)
                    /rpt_family="AluYb" complement(157368..157390)
                    /rpt_family="(TAAAn)" complement(157391..157640)
                    /rpt_family="AluSg" 157641..157687
                    /rpt_family="MR" complement(158254..158275)
                    /rpt_family="(CAAA)n" complement(158288..158555)
                    /rpt_family="AluCo" 158792..159006
                    /rpt_family="AluSg/x" complement(159894..160598)
                    /rpt_family="LIM4" complement(160740..160773)
                    /rpt_family="(CAIA)n" complement(160993..161062)
                    /rpt_family="A-rich" complement(161190..161329)
                    /rpt_family="MR" complement(161706..161780)
                    /rpt_family="L2"

```

```

repeat_region      complement(162242..162312)
                    /rpt_family="MER46A"
repeat_region      /rpt_family="MER2" 162313..162615
                    /rpt_family="MR2" complement(163001..163293)
                    /rpt_family="AluY" 165058..165345
                    /rpt_family="AluJo" complement(16537..165724)
                    /rpt_family="LIMC5" 165939..166032
                    /rpt_family="L2" complement(166033..166330)
                    /rpt_family="AluSx" 166331..166463
                    /rpt_family="L2" complement(166464..166772)
                    /rpt_family="AluSg" 166773..167134
                    /rpt_family="L2" 167654..167869
                    /rpt_family="MR" complement(167932..168222)
                    /rpt_family="AluYb" complement(168264..168439)
                    /rpt_family="MER5A" complement(168508..168669)
                    /rpt_family="MR" complement(168743..169041)
                    /rpt_family="AluSc" 169191..169298
                    /rpt_family="L2"
BASE COUNT 45170 a 35676 c 38297 g 50887 t
ORIGIN

```

```

Query Match      5.7%; Score 35.2; DB 12; Length 170030;
Best Local Similarity 50.0%; Pred. No. 5.8;
Matches 88; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

```

```

QY 448 tgcgtctcgcggagctgcgtccagatttaccggaagaccacactgtgtttcc 507
DB 22332 TGGTGGCCGCTGTAAATCCAGCTACTCAGAGGCTGAGGTGGAATAATGCTTGAACCCA 223391
QY 508 agacatgagatgtgtccctgcggaagaccacactccctcctacgaaccacactgtgtgtg 567
DB 22392 GGAAGGTGGAAGGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 22451
QY 568 atgtgcaagttgcagcagacccctttaaataaagaaataaataaataaataaataaataa 623
DB 22452 CAAGACTCCATCTTAAACACACACACACACACACACACACACACACACACACACACAC 22507

```

```

RESULT 14
AC003829
LOCUS AC005829 170030 bp DNA PRI 15-NOV-1998
DEFINITION Homo sapiens chromosome 17, clone hRPK.259.G.18, complete sequence.
ACCESSION AC005829
NID 93873300
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 170030)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 17, clone hRPK.259.G.18
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 170030)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Brown,A., Castle,A., Cerny,J., Berra,N., Beckely,R., Benn,J., Boutwell,C.,
Collins,S., Collymore,A., Cooke,P., Corliss,D., Depayre,E., Devon,K., Dewar,K.,

```

Donelan, L., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraigery, K., Grant, G., Hagos, B., Heatford, A., Herena, L., Horton, L., Howland, J. C., Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Tichovolsky, N., Torrella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D., Ye, W. J., Zhao, J. and Zody, M.

TITLE
Direct Submission
JOURNAL
Submitted (17-OCT-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE
3 (bases 1 to 170030)
AUTHORS
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M., Baker, J., Baldwin, O., Barna, N., Becker, R., Benn, J., Bouwelle, C., Brown, A., Castle, A., Cerny, J., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Corliss, D., Depayre, E., Devon, K., Dewar, K., Donelan, L., Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D., Gardyna, S., Geraigery, K., Grant, G., Hagos, B., Heatford, A., Herena, L., Horton, L., Howland, J. C., Jacotot, L., Jones, C., Kann, L., Karatas, A., Lehoczy, J., Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Meldrum, J., Molla, M., Morris, W., Morrow, J., Mychaleckyj, J., Nahf, R., Naylor, J., Niloff, M., O'Connor, T., O'Donnell, P., Pavlin, B., Peterson, K., Riley, R., Roberts, D., Roy, A., Severy, P., Stange-Thomann, N., Stillwell, J., Stojanovic, N., Stone, C., Subramanian, A., Tesfaye, S., Tichovolsky, N., Torrella-Miller, I., Vassiliev, H., Vo, A., Wagner, A., Wheeler, J., Wu, Y., Wyman, D., Ye, W. J., Zhao, J. and Zody, M.

TITLE
Direct Submission
JOURNAL
Submitted (15-NOV-1998) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT
On Nov 15, 1998 this sequence version replaced g1:3868742.
All repeats were identified using RepeatMasker: Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Only the first 170030 base pairs of this clone are being submitted.
The remainder overlaps accession number AC004139 (WIGR project 1291).

FEATURES
SOURCE
1. 170030
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="hprk.259.G_18"
/clone_lib="RPC1-11 human BAC library"
/map="17"
/chromosome="17"
2. 271
/rpt_family="L1ME"
476..786
/rpt_family="AluUb"
1750..1915
/rpt_family="L2"
complement(2143..2448)
/rpt_family="AluY"
complement(2600..2810)
/rpt_family="MIR"
complement(5987..6018)
/rpt_family="(CCCTA)n"
6704..6987
/rpt_family="AluSg"
6988..7007
/rpt_family="(TAAA)n"
7141..7173
/rpt_family="AT_rich"
complement(7360..7633)
/rpt_family="rigger3(Golem)"
7634..7942
/rpt_family="AluSx"
complement(7943..8051)
/rpt_family="rigger3(Golem)"

repeat_region complement(8528..8838)
repeat_region /rpt_family="AluSx"
repeat_region complement(8901..9240)
repeat_region /rpt_family="AluUb"
repeat_region 10142..10422
repeat_region /rpt_family="AluSg"
repeat_region 10473..10749
repeat_region /rpt_family="AluUb"
repeat_region complement(10750..10769)
repeat_region /rpt_family="(CAAA)n"
repeat_region complement(10771..11044)
repeat_region /rpt_family="AluUb"
repeat_region 11249..11366
repeat_region /rpt_family="AluSg/X"
repeat_region 11371..11391
repeat_region /rpt_family="AT_rich"
repeat_region 11417..11703
repeat_region /rpt_family="AluUb"
repeat_region 11888..12188
repeat_region /rpt_family="AluSg"
repeat_region 12189..12263
repeat_region /rpt_family="MER41B"
repeat_region 12275..12290
repeat_region /rpt_family="AluUb"
repeat_region 12391..12582
repeat_region /rpt_family="AluSg"
repeat_region 12583..12731
repeat_region /rpt_family="AluUb"
repeat_region 15348..15659
repeat_region /rpt_family="AluYb8"
repeat_region 15721..16004
repeat_region /rpt_family="AluO"
repeat_region 16307..17274
repeat_region /rpt_family="L7R5"
repeat_region 17651..17753
repeat_region /rpt_family="L1MC/D"
repeat_region complement(17754..18051)
repeat_region /rpt_family="AluSc"
repeat_region 18052..18063
repeat_region /rpt_family="L1MC/D"
repeat_region 18167..18481
repeat_region /rpt_family="AluSx"
repeat_region 18486..18796
repeat_region /rpt_family="AluY"
repeat_region 18846..18873
repeat_region /rpt_family="AT_rich"
repeat_region 18945..19239
repeat_region /rpt_family="AluSx"
repeat_region 19243..19509
repeat_region /rpt_family="AluO"
repeat_region complement(19558..19700)
repeat_region /rpt_family="L2"
repeat_region 19709..19968
repeat_region /rpt_family="HAL1"
repeat_region 20013..20318
repeat_region /rpt_family="AluSg"
repeat_region 20343..20369
repeat_region /rpt_family="AT_rich"
repeat_region complement(20370..20495)
repeat_region /rpt_family="FLAM_C"
repeat_region complement(20502..20797)
repeat_region /rpt_family="AluSg"
repeat_region 20798..20815
repeat_region /rpt_family="AT_rich"
repeat_region 20866..21093
repeat_region /rpt_family="HAL1"
repeat_region 21094..21398
repeat_region /rpt_family="AluUb"
repeat_region 21405..21715
repeat_region /rpt_family="AluSx"
repeat_region 21716..21785
repeat_region /rpt_family="HAL1"
repeat_region 21791..22093

```
repeat_region      /rpt_family="AluSg"      22187. .22465
                    /rpt_family="AluSg1"    22466. .22521
repeat_region      /rpt_family="C(AAA)n"    22643. .22938
repeat_region      /rpt_family="AluSx"     22947. .23251
repeat_region      /rpt_family="AluSg"     23784. .24097
repeat_region      /rpt_family="AluUb"     24698. .24993
repeat_region      /rpt_family="AluSg"     25122. .25147
repeat_region      /rpt_family="(CA)n"     25698. .26013
repeat_region      /rpt_family="AluSp"     26078. .26104
repeat_region      /rpt_family="AT_rich"    complement(27061. .27091)
repeat_region      /rpt_family="C(AAA)n"    28173. .28481
repeat_region      /rpt_family="AluSg"     28829. .29117
repeat_region      /rpt_family="AluSg"     29407. .29727
repeat_region      /rpt_family="AluSp"     complement(29974. .30261)
repeat_region      /rpt_family="AluSg"     31297. .31320
repeat_region      /rpt_family="AT_rich"    complement(31353. .31636)
repeat_region      /rpt_family="AluSp"     31885. .31907
repeat_region      /rpt_family="AT_rich"    32571. .32718
repeat_region      /rpt_family="AluSx"     32824. .32994
repeat_region      /rpt_family="AluSx"     33705. .33731
repeat_region      /rpt_family="GC_rich"    complement(33856. .33971)
repeat_region      /rpt_family="(CCGC)n"    34016. .34043
repeat_region      /rpt_family="GC_rich"    complement(34842. .35072)
repeat_region      /rpt_family="AluSx"     35111. .35419
repeat_region      /rpt_family="AluSg"     35420. .35517
repeat_region      /rpt_family="purine-rich" 35518. .35638
repeat_region      /rpt_family="(GAA)n"    complement(36572. .36603)
repeat_region      /rpt_family="L2"        36626. .36741
repeat_region      /rpt_family="FLAM_C"    36779. .36821
repeat_region      /rpt_family="MER77"    complement(36849. .36877)
repeat_region      /rpt_family="(TAAA)n"    36892. .37008
repeat_region      /rpt_family="AluSg/x"    complement(37014. .37315)
repeat_region      /rpt_family="AluSp"     complement(37316. .37618)
repeat_region      /rpt_family="AluSx"     complement(37643. .37671)
repeat_region      /rpt_family="C(AAA)n"    complement(37672. .37955)
repeat_region      /rpt_family="AluUb"     complement(38073. .38228)
repeat_region      /rpt_family="AluSp/q"

repeat_region      38377. .38500
                    /rpt_family="FLAM_C"    38501. .38528
repeat_region      /rpt_family="(GAA)n"    38529. .38630
repeat_region      /rpt_family="FLAM_C"    38639. .38926
repeat_region      /rpt_family="AluSx"     39001. .39036
repeat_region      /rpt_family="AT_rich"    complement(39115. .39152)
repeat_region      /rpt_family="(GAA)n"    39176. .39471
repeat_region      /rpt_family="AluSx"     39486. .39644
repeat_region      /rpt_family="MLTIF"    complement(39673. .39963)
repeat_region      /rpt_family="LIPB1"    39970. .40272
repeat_region      /rpt_family="AluSp"     41835. .41868
repeat_region      /rpt_family="POLY_A"    41871. .42183
repeat_region      /rpt_family="AluY"     complement(42184. .42379)
repeat_region      /rpt_family="LIPB1"    42380. .42671
repeat_region      /rpt_family="AluSc"    complement(42672. .42691)
repeat_region      /rpt_family="LIPB1"    42754. .43049
repeat_region      /rpt_family="MLTIF"    complement(43054. .43356)
repeat_region      /rpt_family="AluUb"    43363. .43393
repeat_region      /rpt_family="AT_rich"    complement(43400. .43414)
repeat_region      /rpt_family="MIR"     complement(43415. .43723)
repeat_region      /rpt_family="AluSx"    complement(43724. .43884)
repeat_region      /rpt_family="MIR"     complement(44278. .44464)
repeat_region      /rpt_family="MLT1A1"  complement(44465. .44774)
repeat_region      /rpt_family="AluSx"    complement(44775. .44909)
repeat_region      /rpt_family="MLT1A1"  44914. .45213
repeat_region      /rpt_family="AluY"     45223. .45527
repeat_region      /rpt_family="AluSp"    45620. .45752
repeat_region      /rpt_family="AluUb"    45753. .46028
repeat_region      /rpt_family="AluSx"    46031. .46055
repeat_region      /rpt_family="(C(AAA)n"  46056. .46342
repeat_region      /rpt_family="AluSg"    46343. .46421
repeat_region      /rpt_family="AluUb"    46451. .46500
repeat_region      /rpt_family="(TAAA)n"  complement(46504. .46697)
repeat_region      /rpt_family="HRRVL"    complement(46712. .46811)
repeat_region      /rpt_family="MERVL"    complement(47100. .48514)
repeat_region      /rpt_family="SVA"      48524. .48673
repeat_region      /rpt_family="AluSg/x"    complement(49245. .49330)
```

```

repeat_region      /rpt_family="MER20"
                    49331..49629
repeat_region      /rpt_family="AluSp"
                    complement(49630..49711)
repeat_region      /rpt_family="MER20"
                    complement(49773..49870)
repeat_region      /rpt_family="AluIo/FRAM"
                    49952..50063
repeat_region      /rpt_family="Purine-rich"
                    complement(50064..50135)
repeat_region      /rpt_family="AluIo/FLAM"
                    complement(50136..50563)
repeat_region      /rpt_family="MLT2A"
                    50658..50789
repeat_region      /rpt_family="AluSg/x"
                    50792..51074
repeat_region      /rpt_family="AluSg"
                    complement(51313..51533)
repeat_region      /rpt_family="L2"
                    complement(51552..51673)
repeat_region      /rpt_family="AluSg/x"
                    51674..51713
repeat_region      /rpt_family="L1ME1"
                    51714..52008
repeat_region      /rpt_family="MER2"
                    52009..52068
repeat_region      /rpt_family="L1ME1"
                    52069..52361
repeat_region      /rpt_family="AluSx"
                    52362..52652
repeat_region      /rpt_family="L1ME1"
                    complement(53828..53857)
repeat_region      /rpt_family="(CAAAA)n"
                    complement(53858..54129)
repeat_region      /rpt_family="AluSx"
                    54339..54484
repeat_region      /rpt_family="L2"
                    complement(54679..54876)
repeat_region      /rpt_family="MIR"
                    complement(54955..55100)
repeat_region      /rpt_family="polypurine"
                    57139..57172
repeat_region      /rpt_family="GC-rich"
                    57259..57279
repeat_region      /rpt_family="GC-rich"
                    complement(57293..57683)
repeat_region      /rpt_family="Purine-rich"
                    complement(57790..57927)
repeat_region      /rpt_family="L2"
                    complement(58042..58151)
repeat_region      /rpt_family="MER21B"
                    58233..58473
repeat_region      /rpt_family="AluIo"
                    58474..58752
repeat_region      /rpt_family="AluSg1"
                    complement(58929..59240)
repeat_region      /rpt_family="AluIo"
                    complement(59340..59645)
repeat_region      /rpt_family="AluSp"
                    complement(59652..59791)
repeat_region      /rpt_family="FLAM.C"
                    complement(59794..59841)
repeat_region      /rpt_family="(TAAA)n"
                    complement(59842..60119)
repeat_region      /rpt_family="AluSx"
                    60120..60145
repeat_region      /rpt_family="AT-rich"
                    complement(60269..60398)
repeat_region      /rpt_family="MLT1A2"
                    complement(60399..60502)
repeat_region      /rpt_family="MER6B"
                    complement(60503..60788)
repeat_region      /rpt_family="AluSx"

repeat_region      complement(60789..61024)
repeat_region      /rpt_family="MER6B"
                    61025..61312
repeat_region      /rpt_family="AluSg"
                    complement(61313..61485)
repeat_region      /rpt_family="MER6B"
                    complement(61486..61639)
repeat_region      /rpt_family="MLT1A2"
                    complement(61647..61815)
repeat_region      /rpt_family="AluSx"
                    complement(61816..62123)
repeat_region      /rpt_family="AluYb8"
                    complement(62124..62275)
repeat_region      /rpt_family="AluSx"
                    62276..62545
repeat_region      /rpt_family="L1M4"
                    62571..62662
repeat_region      /rpt_family="AluSg"
                    complement(62663..63313)
repeat_region      /rpt_family="SVA"
                    63314..63369
repeat_region      /rpt_family="AluSg"
                    63370..63601
repeat_region      /rpt_family="AluSg"
                    63604..63708
repeat_region      /rpt_family="L1M4"
                    complement(63785..64085)
repeat_region      /rpt_family="AluY"
                    64101..64161
repeat_region      /rpt_family="L1M4"
                    complement(64162..64333)
repeat_region      /rpt_family="FRAM"
                    64334..64837
repeat_region      /rpt_family="L1M4"
                    complement(64838..64988)
repeat_region      /rpt_family="AluIo"
                    complement(64989..65282)
repeat_region      /rpt_family="AluSp"
                    complement(65283..65430)
repeat_region      /rpt_family="AluIo"
                    65431..65646
repeat_region      /rpt_family="L1M4"
                    complement(65647..65815)
repeat_region      /rpt_family="AluSc"
                    complement(65816..66116)
repeat_region      /rpt_family="AluSp"
                    complement(66117..66161)
repeat_region      /rpt_family="AluSc"
                    66162..66229
repeat_region      /rpt_family="L1M4"
                    66232..66302
repeat_region      /rpt_family="L1ME"
                    66325..66368
repeat_region      /rpt_family="AluSg/x"
                    66369..66671
repeat_region      /rpt_family="AluSx"
                    66672..66852
repeat_region      /rpt_family="AluSg/x"
                    66909..66915
repeat_region      /note="Single-stranded coverage."
                    66912..67000
repeat_region      /rpt_family="HAL1"
                    67001..67305
repeat_region      /rpt_family="AluSx"
                    67306..67363
repeat_region      /rpt_family="HAL1"
                    complement(67372..67530)
repeat_region      /rpt_family="FRAM"
                    67539..67688
repeat_region      /rpt_family="AluIo/FLAM"
                    complement(67689..68034)
repeat_region      /rpt_family="AluSx"
                    complement(68036..68328)

```

```
repeat_region /rpt_family="AluSg" complement(68427..68721)
repeat_region /rpt_family="AluSg" 68750..69186
repeat_region /rpt_family="L2" 69247..69346
repeat_region /rpt_family="MERSA" complement(69347..69551)
repeat_region /rpt_family="AluSg" 69552..69859
repeat_region /rpt_family="AluY" complement(69860..69985)
repeat_region /rpt_family="AluSg" 69986..70026
repeat_region /rpt_family="MERSA" 70172..70287
repeat_region /rpt_family="L2" complement(70291..70564)
repeat_region /rpt_family="AluYo" 70566..70626
repeat_region /rpt_family="Alu" 70690..70771
repeat_region /rpt_family="HALI" 70779..70807
repeat_region /rpt_family="Alu" 71156..71189
repeat_region /rpt_family="(TA)n" complement(71249..71356)
repeat_region /rpt_family="MIR" complement(71550..71586)
repeat_region /rpt_family="tRNA-Tyr-TTC" complement(71616..71906)
repeat_region /rpt_family="AluYo" 72074..72106
repeat_region /rpt_family="L2" 72174..72205
repeat_region /rpt_family="MER4D" 72212..72233
repeat_region /rpt_family="AT_rich" complement(72310..72609)
repeat_region /rpt_family="AluSg" complement(72863..72880)
repeat_region /rpt_family="AluYb" complement(72881..72917)
repeat_region /rpt_family="(CAA)n" complement(72918..73202)
repeat_region /rpt_family="AluYb" 73232..73537
repeat_region /rpt_family="AluY" complement(73542..73635)
repeat_region /rpt_family="MIR" complement(73717..73994)
repeat_region /rpt_family="AluSg" complement(74083..74393)
repeat_region /rpt_family="AluSx" 74529..74717
repeat_region /rpt_family="L2" 75055..75100
repeat_region /rpt_family="AT_rich" complement(75524..75896)
repeat_region /rpt_family="LIM3e" 76331..76355
repeat_region /rpt_family="AT_rich" 78004..78036
repeat_region /rpt_family="AT_rich" 79901..80186
repeat_region /rpt_family="AluSx" 80340..80372
repeat_region /rpt_family="AT_rich" 80373..80668
repeat_region /rpt_family="AluYb8" complement(80759..81066)
repeat_region /rpt_family="AluY"

repeat_region 81223..81533
repeat_region /rpt_family="AluSg" complement(81615..81909)
repeat_region /rpt_family="AluSc" 81911..82058
repeat_region /rpt_family="MIR" complement(82092..82382)
repeat_region /rpt_family="AluSx" 82445..82612
repeat_region /rpt_family="MIR" complement(82641..82805)
repeat_region /rpt_family="MIR1B" complement(82848..83145)
repeat_region /rpt_family="AluSx" complement(83146..83281)
repeat_region /rpt_family="AluYb" complement(83282..83404)
repeat_region /rpt_family="MIR1B" complement(83405..83694)
repeat_region /rpt_family="AluSx" complement(83695..83756)
repeat_region /rpt_family="MIR1B" complement(83757..84065)
repeat_region /rpt_family="AluYo" complement(84066..84114)
repeat_region /rpt_family="MIR1B" complement(84272..84557)
repeat_region /rpt_family="AluSg" 84563..84818
repeat_region /rpt_family="L2" complement(84819..85129)
repeat_region /rpt_family="AluSx" 85130..85480
repeat_region /rpt_family="L2" 85717..85816
repeat_region /rpt_family="HY3" complement(86213..86516)
repeat_region /rpt_family="AluSc" complement(86599..87253)
repeat_region /rpt_family="AluSx" 87281..87573
repeat_region /rpt_family="AluYb8" 87574..87599
repeat_region /rpt_family="(CAA)n" 87600..87612
repeat_region /rpt_family="AluYb8" 87891..88354
repeat_region /rpt_family="LIMC/D" 88370..88663
repeat_region /rpt_family="AluSg" complement(89376..89597)
repeat_region /rpt_family="MER30" complement(90385..90696)
repeat_region /rpt_family="AluSg" 91095..91396
repeat_region /rpt_family="AluSx" complement(91689..91997)
repeat_region /rpt_family="AluSg" complement(92020..92131)
repeat_region /rpt_family="AluSg/X" complement(92191..92249)
repeat_region /rpt_family="MIR1A2" complement(92251..92558)
repeat_region /rpt_family="AluYo" complement(92745..93059)
repeat_region /rpt_family="AluSg" complement(93060..93192)
repeat_region /rpt_family="AluYb" complement(93193..93293)
repeat_region /rpt_family="LIM4" complement(93294..93353)
repeat_region /rpt_family="AluYo" complement(93354..93658)
```

```
repeat_region /rpt_family="AluSg1" complement(93859. .93887)
repeat_region /rpt_family="AluJo" complement(93888. .94306)
repeat_region /rpt_family="L1M4" complement(94312. .94443)
repeat_region /rpt_family="FLAM_A" complement(94445. .94777)
repeat_region /rpt_family="L1M4" complement(94778. .95092)
repeat_region /rpt_family="AluSg" complement(95093. .95189)
repeat_region /rpt_family="L1M4" complement(95190. .95488)
repeat_region /rpt_family="AluSx" complement(95489. .96317)
repeat_region /rpt_family="L1M4" complement(96413. .96713)
repeat_region /rpt_family="AluSx" complement(96714. .96992)
repeat_region /rpt_family="AluJo" complement(97632. .97938)
repeat_region /rpt_family="AluSg" complement(98515. .98684)
repeat_region /rpt_family="MER46B" complement(98685. .98866)
repeat_region /rpt_family="AluJo" complement(98867. .99159)
repeat_region /rpt_family="AluCb" complement(99160. .99298)
repeat_region /rpt_family="AluJo" complement(99299. .99480)
repeat_region /rpt_family="MER46B" 99677. 99878
repeat_region /rpt_family="AluSx" 100115. 100142
repeat_region /rpt_family="AT_rich" complement(100143. .100427)
repeat_region /rpt_family="AluY" 100428. 100445
repeat_region /rpt_family="AT_rich" complement(100822. .100849)
repeat_region /rpt_family="(CA)n" complement(102147. .102436)
repeat_region /rpt_family="AluSx" complement(102439. .102465)
repeat_region /rpt_family="POLY_A" complement(102608. .102751)
repeat_region /rpt_family="MIR" complement(102756. .102859)
repeat_region /rpt_family="L2" complement(103336. .10357)
repeat_region /rpt_family="AluSx" complement(103725. .103756)
repeat_region /rpt_family="polypurine" complement(103757. .104063)
repeat_region /rpt_family="AluJo" complement(104625. .104928)
repeat_region /rpt_family="AluCb" 105057. 105317
repeat_region /rpt_family="AluJo" 106712. 106734
repeat_region /rpt_family="GC_rich" 107291. 107369
repeat_region /rpt_family="GC_rich" complement(107433. .107457)
repeat_region /rpt_family="(CGAC)n" 107595. 107618
repeat_region /rpt_family="GC_rich" 107733. 107762
repeat_region /rpt_family="GC_rich" 107928. 108001
repeat_region /rpt_family="GC_rich"

repeat_region complement(108073. 108129)
repeat_region /rpt_family="purine-rich" complement(110962. .111287)
repeat_region /rpt_family="AluSx" complement(111508. .111578)
repeat_region /rpt_family="L1M3e" complement(111639. .111804)
repeat_region /rpt_family="FAM" complement(111827. .112139)
repeat_region /rpt_family="AluSx" 112907. 112928
repeat_region /rpt_family="AT_rich" 113857. 114160
repeat_region /rpt_family="AluSx" complement(114487. .114794)
repeat_region /rpt_family="AluYb8" complement(115026. .115147)
repeat_region /rpt_family="FLAM_C" complement(115154. .115233)
repeat_region /rpt_family="(CA)n" complement(115676. .115761)
repeat_region /rpt_family="L2" 115762. 116050
repeat_region /rpt_family="AluSg" complement(116051. .116264)
repeat_region /rpt_family="L2" 11638. 116700
repeat_region /rpt_family="L2" complement(116960. .117219)
repeat_region /rpt_family="AluSx" complement(117220. .117335)
repeat_region /rpt_family="GC_rich" complement(11736. .117590)
repeat_region /rpt_family="AluYb" complement(117632. .117741)
repeat_region /rpt_family="FLAM_A" 117742. 117809
repeat_region /rpt_family="L2" 118052. 118266
repeat_region /rpt_family="MIR" complement(118930. .119080)
repeat_region /rpt_family="L2" 119157. 119454
repeat_region /rpt_family="AluSg" complement(119486. .119609)
repeat_region /rpt_family="L2" 119798. 119961
repeat_region /rpt_family="L2" 119993. 120040
repeat_region /rpt_family="(TA)n" complement(120042. .120170)
repeat_region /rpt_family="FLAM_C" complement(120180. .120356)
repeat_region /rpt_family="AluSg/*" complement(120360. .120674)
repeat_region /rpt_family="AluSg" complement(120809. .121101)
repeat_region /rpt_family="AluSg" 121680. 121893
repeat_region /rpt_family="MER20" 122812. 123110
repeat_region /rpt_family="AluYb8" complement(123636. .123802)
repeat_region /rpt_family="MIR" complement(124180. .124479)
repeat_region /rpt_family="AluSg" 125035. 125070
repeat_region /rpt_family="AT_rich" 125239. 125524
repeat_region /rpt_family="AluSx" 125525. 125577
repeat_region /rpt_family="(CA)n" complement(125807. .125875)
```

```
/rpt_family="Charl1el"
repeat_region 125903. .125924
/rpt_family="AT_r1ch"
repeat_region 125933. .126199
/rpt_family="AluUb"
repeat_region complement(126217. .126407)
/rpt_family="Char1le1a"
repeat_region 126408. .126681
/rpt_family="AluSc"
repeat_region complement(126682. .127379)
/rpt_family="Char1le1a"
repeat_region complement(127634. .128101)
/rpt_family="L2"
repeat_region complement(130705. .130741)
/rpt_family="(CAAA)n"
repeat_region complement(130746. .130989)
/rpt_family="AluSg"
repeat_region 132238. .132536
/rpt_family="AluSp"
repeat_region complement(133820. .133882)
/rpt_family="A_r1ch"
repeat_region complement(134361. .134658)
/rpt_family="AluSg"
repeat_region complement(134660. .134959)
/rpt_family="AluSx"
repeat_region 135078. .135375
/rpt_family="AluSx"
repeat_region 135378. .135669
/rpt_family="AluY"
repeat_region complement(135741. .135752)
/rpt_family="AluSg/x"
repeat_region complement(135753. .135779)
/rpt_family="(CAT)n"
repeat_region complement(135780. .135944)
/rpt_family="AluSg/x"
repeat_region 136851. .136881
/rpt_family="MIR"
repeat_region 137363. .137522
/rpt_family="FRAM"
repeat_region complement(137618. .138412)
/rpt_family="L2"
repeat_region 138422. .138444
/rpt_family="AT_r1ch"
repeat_region complement(139440. .139755)
/rpt_family="AluY"
repeat_region complement(139820. .140012)
/rpt_family="MIR"
repeat_region 140289. .140579
/rpt_family="AluUb"
repeat_region complement(140756. .140818)
/rpt_family="MIR"
repeat_region 142048. .142358
/rpt_family="AluSg"
repeat_region 142359. .142524
/rpt_family="AluSg/x"
repeat_region complement(143972. .144135)
/rpt_family="AluSg/x"
repeat_region 144187. .144263
/rpt_family="FRAM"
repeat_region 144276. .144298
/rpt_family="AT_r1ch"
repeat_region complement(144535. .144820)
/rpt_family="AluSx"
repeat_region complement(144945. .145101)
/rpt_family="FRAM"
repeat_region complement(145210. .145403)
/rpt_family="MER20"
repeat_region 145431. .145732
/rpt_family="AluSp"
repeat_region 145934. .146027
/rpt_family="FRAM_C"
repeat_region 146309. .146538
/rpt_family="AluUb"
repeat_region 148821. .148988
/rpt_family="MER5B"
repeat_region complement(149198. .149678)
/rpt_family="L2"
repeat_region complement(150807. .150838)
/rpt_family="AluYb8"
repeat_region complement(150839. .150864)
/rpt_family="(TAA)n"
repeat_region complement(150865. .151162)
/rpt_family="AluYb8"
repeat_region 151247. .151464
/rpt_family="MER58A"
repeat_region 151588. .151891
/rpt_family="AluY"
repeat_region 151970. .151990
/rpt_family="AT_r1ch"
repeat_region 152003. .152297
/rpt_family="AluSx"
repeat_region complement(152753. .152937)
/rpt_family="MIR"
repeat_region complement(153354. .153291)
/rpt_family="(GA)n"
repeat_region complement(153392. .153595)
/rpt_family="AluUb"
repeat_region 153812. .153834
/rpt_family="AT_r1ch"
repeat_region complement(153843. .153965)
/rpt_family="FRAM_A"
repeat_region complement(154028. .154336)
/rpt_family="AluYas"
repeat_region 154373. .154415
/rpt_family="MER6B"
repeat_region 154461. .154624
/rpt_family="MIR"
repeat_region 154864. .155159
/rpt_family="AluSx"
repeat_region 156920. .157056
/rpt_family="MIR"
repeat_region complement(157057. .157365)
/rpt_family="AluYb8"
repeat_region complement(157368. .157390)
/rpt_family="(TAA)n"
repeat_region complement(157391. .157640)
/rpt_family="AluSg"
repeat_region 157641. .157687
/rpt_family="MIR"
repeat_region complement(158254. .158275)
/rpt_family="(CAAA)n"
repeat_region complement(158288. .158555)
/rpt_family="AluUb"
repeat_region 158792. .159006
/rpt_family="AluSg/x"
repeat_region complement(159894. .160598)
/rpt_family="LIM4"
repeat_region complement(160740. .160773)
/rpt_family="(CAT)n"
repeat_region complement(160993. .161062)
/rpt_family="A_r1ch"
repeat_region complement(161190. .161329)
/rpt_family="MIR"
repeat_region complement(161706. .161780)
/rpt_family="L2"
repeat_region complement(162242. .162312)
/rpt_family="MER46A"
repeat_region 162313. .162615
/rpt_family="MER2"
repeat_region complement(163001. .163293)
/rpt_family="AluY"
repeat_region 165058. .165345
/rpt_family="AluUb"
repeat_region complement(165537. .165724)
/rpt_family="LIMC5"
repeat_region 165939. .166032
```

```
repeat_region /rpt_family="L2"
complement(166033..166330)
/rpt_family="AluSx"
repeat_region 166331..166463
/rpt_family="L2"
repeat_region complement(166464..166772)
/rpt_family="AluSg"
repeat_region 166773..167134
/rpt_family="L2"
repeat_region 167135..167654
/rpt_family="L2"
repeat_region 167655..167869
/rpt_family="MIR"
repeat_region complement(167932..168222)
/rpt_family="AluJb"
repeat_region complement(168264..168439)
/rpt_family="MER5A"
repeat_region complement(168508..168669)
/rpt_family="MIR"
repeat_region complement(168743..169041)
/rpt_family="AluSc"
repeat_region 169191..169298
/rpt_family="L2"
BASE COUNT 45170 a 35676 c 38297 g 50887 t
ORIGIN

Query Match 5.7%; Score 35.2; DB 42; Length 170030;
Best Local Similarity 50.0%; Pred. No. 5.8;
Matches 88; Conservative 0; Mismatches 88; Indels 0; Gaps 0;

QY 448 tgcgtctcgaggagacgctcgcagatttaccggaagacaaactggtgtttca 507
Db 22332 TGGTAGCGCCTGTATCCAGTCTCAGAGGCTGAGTGAATAATCGCTTGACCCA 22391

QY 508 agacatggaagtggtgcccctcgagagaccactctcctcgaaccacactgtgagt 567
Db 22392 GGAGGTGAGGTGGAGTGCATGATCACACCACTGCTAGCTGCTGGGTGACAGAG 22451

QY 568 atgtcccaagtgcgcagacccttaaaaaaaaaaagacaaacaaac 623
Db 22452 CAAGCTCCTCATCTTAACACACACAAAAAATTAACACACAGAAC 22507

RESULT 15
AGU79997 2055 bp DNA PLN 06-MAR-1998
LOCUS Atriplex garrettii male-specific DNA sequence.
DEFINITION Atriplex garrettii male-specific DNA sequence.
ACCESSION U79997
NID g2160708
KEYWORDS
SOURCE Atriplex garrettii.
ORGANISM Atriplex garrettii.
Eukaryotes; mitochondrial eukaryotes; Viridiplantae;
Charophyta/Embryophyta group; Embryophyta; Magnoliophyta;
Magnoliopsida; Caryophyllales; Chenopodiales; Atriplex.
1 (bases 1 to 2055)
Ruas,C.F., Fairbanks,D.J., Evans,R.P., Stutz,H.C., Andersen,W.R.
and Ruas,P.M.
Male-specific DNA in Atriplex garrettii
Am. J. Bot. 85, 162-167 (1998)
REFERENCE
AUTHORS Ruas,C.F., Fairbanks,D.J., Evans,R.P., Stutz,H.C., Andersen,W.R.
and Ruas,P.M.
TITLE Direct Submission
JOURNAL Submitted (26-NOV-1996) Botany and Range Science, Brigham Young
UNIVERSITY, 401 WIDB, Provo, UT 84602, USA
FEATURES
location/Qualifiers
1..2055
/organism="Atriplex garrettii"
/db_xref="taxon:55308"
1..2055
/misc_feature
/feature="This sequence was found in all male plants examined
and was absent in all female plants examined."
555..648
misc_feature
```

```
/note="similar to the Antirrhinum majus (snapdragon) CEN
gene product encoded by GenBank Accession Number S81193
and the Arabidopsis thaliana TFL1 gene product encoded by
GenBank Accession Number U77674; possible pseudogene; the
reading frame is interrupted by repetitive DNA 31 codons
upstream from the stop codon at the point where similarity
to the CEN and TFL1 genes is lost"
BASE COUNT 708 a 311 c 338 g 698 t
ORIGIN

Query Match 5.6%; Score 35; DB 8; Length 2055;
Best Local Similarity 61.5%; Pred. No. 4;
Matches 56; Conservative 0; Mismatches 35; Indels 0; Gaps 0;

QY 532 agaccacctcctccctcagacacacactgtggtgagtgccagtgagcagacctt 591
Db 1681 ATACATATATTCATCAATATATCCATATATCTAGGTCTTACCTAAGTCAATATCTT 1622

QY 592 taataaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaaagaa 622
Db 1621 CAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1591
```

Search completed: May 14, 1999, 03:35:05
Job time: 5724 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 11:42:30 ; Search time 129.53 Seconds

(without alignments)
904.854 Million cell updates/sec

Title: US-09-212-270-9

Sequence: 1 gaggcagaattcgcacgag.....aaaaaaaaacaacaac 623

Scoring table: IDENTITY_NUC

Searched: 240622 seqs, 94065609 residues

Database : N_Geneseq_34:*

Pred. No. Is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

| Result | No. | Score | Query | Length | DB | ID | Description |
|--------|-----|-------|-------|--------|----|--------|---------------------|
| | 1 | 326.6 | 52.4 | 1704 | 1 | V33361 | Nucleotide sequenc |
| | 2 | 326.6 | 52.4 | 1496 | 1 | V33362 | Nucleotide sequenc |
| | 3 | 248.8 | 39.9 | 371 | 1 | V11422 | Human secreted pro |
| | 4 | 70.2 | 11.3 | 181 | 1 | V11423 | Human secreted pro |
| | 5 | 35 | 5.6 | 1912 | 1 | Q10388 | Encodes skin-speci |
| | 6 | 35 | 5.6 | 1913 | 1 | T05202 | Human secreted pro |
| | 7 | 34.2 | 5.5 | 1013 | 1 | V39359 | Murine retinoic ac |
| | 8 | 34.2 | 5.5 | 1013 | 1 | V39359 | Humanised anti-HM |
| | 9 | 33.8 | 5.4 | 2634 | 1 | V05079 | Protein recognisese |
| | 10 | 33.4 | 5.4 | 1016 | 1 | V59114 | Bax inhibitor Bi-1 |
| | 11 | 33.2 | 5.3 | 3692 | 1 | T72107 | Nucleotide sequenc |
| | 12 | 33 | 5.3 | 2083 | 1 | T89343 | Rat semaphorin 2.9 |
| | 13 | 32.8 | 5.3 | 1868 | 1 | T73502 | Human p53 CDNA #1. |
| | 14 | 32.6 | 5.2 | 1227 | 1 | Q54832 | Poplar 1-aminocycl |
| | 15 | 32.6 | 5.2 | 1227 | 1 | T37306 | Sequence of clone |
| | 16 | 32.4 | 5.2 | 1189 | 1 | V41452 | Murine gonadotropi |
| | 17 | 32.4 | 5.2 | 1376 | 1 | V34157 | Nucleotide sequenc |
| | 18 | 32.2 | 5.2 | 1621 | 1 | V58584 | Human secreted pro |
| | 19 | 32.2 | 5.2 | 1621 | 1 | V61199 | Prostate tumour sp |
| | 20 | 32 | 5.1 | 2481 | 1 | T87940 | Full length CDNA s |
| | 21 | 32 | 5.1 | 2604 | 1 | T73595 | Rat neuropeptide y |
| | 22 | 31.8 | 5.1 | 1964 | 1 | N81586 | Rat neuropeptide y |
| | 23 | 31.8 | 5.1 | 1964 | 1 | N92781 | CBF-20 CDNA encod |
| | 24 | 31.8 | 5.1 | 1964 | 1 | O54354 | CDNA sequence enc |
| | 25 | 31.8 | 5.1 | 1338 | 1 | Q65607 | Carthoembryonic a |
| | 26 | 31.8 | 5.1 | 1964 | 1 | T46064 | Rabbit zona pelluc |
| | 27 | 31.8 | 5.1 | 5873 | 1 | T61701 | Carthoembryonic a |
| | 28 | 31.8 | 5.1 | 1338 | 1 | T64789 | Human prostatic sp |
| | 29 | 31.6 | 5.1 | 1174 | 1 | T39050 | Rabbit ZPC CDNA. I |
| | 30 | 31.6 | 5.1 | 1386 | 1 | V11835 | CDNA encoding cell |
| | 31 | 31.6 | 5.1 | 1174 | 1 | V39096 | Homo sapiens tub I |
| | 32 | 31.6 | 5.1 | 712 | 1 | V59811 | Monoclonal endo |
| | 33 | 31.6 | 5.1 | 907 | 1 | V59685 | Human secreted pro |
| | 34 | 31.4 | 5.0 | 755 | 1 | Q03970 | Human secreted pro |
| | 35 | 31.4 | 5.0 | 737 | 1 | V34270 | Clone p25b New pr |
| | 36 | 31.4 | 5.0 | 688 | 1 | V34216 | Human secreted pro |
| | 37 | 31 | 5.0 | 1226 | 1 | O24919 | Human secreted pro |
| | 38 | 31 | 5.0 | 1374 | 1 | O68434 | Human interferon-g |
| | 39 | 31 | 5.0 | 971 | 1 | Q73760 | Human hml CDNA. Hml |
| | 40 | 31 | 5.0 | 971 | 1 | T35704 | RP-8 Programmed ce |
| | 41 | 31 | 5.0 | 966 | 1 | T87430 | Coding sequence f |
| | 42 | 31 | 5.0 | 1040 | 1 | V02818 | Clone H174-10. Poi |
| | 43 | 31 | 5.0 | 3114 | 1 | V02800 | Human R18AM genom |
| | | | | | | | Human R18AM CDNA. |

| | | | | | |
|----|----|-----|---|--------|--------------------|
| 44 | 31 | 5.0 | 1 | T90169 | Oil seed rape cyst |
| 45 | 31 | 5.0 | 1 | V10120 | Human retinoid rec |

ALIGNMENTS

| | |
|-----------------------|---|
| RESULT | 1 |
| V33361 | |
| ID | V33361 standard; cDNA to mRNA; 1704 BP. |
| AC | V33361. |
| DT | 02-DEC-1998 (first entry) |
| DW | Nucleotide sequence of human alpha-OAF065. |
| KW | Human; alpha-OAF065; stroma cell; antibody; inflammatory; |
| KW | cytokine-mediated disease; rheumatism; ulcerative colitis; ss. |
| OS | Homo sapiens. |
| FH | Key |
| FT | CDS |
| FT | Location/Qualifiers |
| FT | /tag= a |
| FT | /product= "human alpha-OAF065 protein" |
| FT | sig_peptide |
| FT | 45..119 |
| FT | /*tag= b |
| FT | mat_peptide |
| FT | 120..1295 |
| FT | /*tag= C |
| FT | /transl_except= (pos:711..713, aa= Pro) |
| FT | /transl_except= (pos:714..716, aa= Arg) |
| PN | MO9838304-A1. |
| PD | 03-SEP-1998. |
| PF | 26-FEB-1998; J00799. |
| PR | 27-FEB-1997; JP-043143. |
| PA | (ONOY) ONO PHARM CO LTD. |
| P1 | Fukushima D, Konishi M, Tada H; |
| DR | WP1; 98-481205/41. |
| P | P-PSDB; W70386. |
| PT | Membrane polypeptide expressed by human stroma cells, and antibodies |
| PT | recognising it for treatment of inflammatory and other |
| PT | cytokine-mediated diseases. |
| PS | Claim 5; Pages 31-32; 54pp; Japanese. |
| CC | This is the nucleotide sequence of the human alpha-OAF065, used in |
| CC | the method of the invention. The process involves the use of peptides |
| CC | expressed by stroma cells, and its antibodies are used for in the |
| CC | prevention and treatment of inflammatory and other cytokine-mediated |
| CC | diseases such as rheumatism, ulcerative colitis. |
| SQ | Sequence I/04 BP; 429 A; 426 C; 430 G; 419 T; |
| Query Match: | 52.4%; Score 326.6; DB 1; Length 1704; |
| Best Local Similarity | 78.2%; Pred. No. 2.5e+86; |
| Matches 392; | Conservative 0; Mismatches 109; Indels 0; Gaps 0. |
| OY | 89 caggaataaacacgcttgcgtgagagccatcggcatcaagaagtctactacacagagacg 148 |
| Db | 18 CAACAATAAATTCATTGTGATAAGAAAGATGGCTTAAAGTGCTACTACAGAACAAGAAA 77 |
| OY | 149 gtgccttcgcgcaccattctctctactctcacccctcgcatgaaagtgtgtcgaaaac 208 |
| Db | 78 ACGTTTTCACCTCTTGTATGTATTACTAGGCTATTTGTCAATGAAGTGACTGTGAAAAA 137 |
| OY | 209 ggaagattcgaagcagcagaagaattcaagaagcatcgatctggaactgtgctcttcgaacag 268 |
| Db | 138 GGAGACTGTAGACACCAAGAAATTCAGGGATCGGTGTGAAATGTGTGCCGCAACACAG 197 |
| OY | 269 tgcgaccttgcatggaattgttccaaggaaatgtgcttcgtgctaagggagagatgcacag 328 |
| Db | 198 TGTGGGCCAAGGAGTGAATTTGCTTAAGGAAATGTGCCTTCGCTATGGGAGGATGCACAG 257 |
| OY | 329 ttgtgcccttcgaagccgacccggtgtcaagaagaaagctgggtttccacaagtgtaagcca 388 |
| Db | 258 TGTGTGAGACTGCCGCTTCACAGATTCCAAGAGAGACTGGGCTTCAGAAATTCAGAACCC 317 |
| OY | 389 ttgtcgacctgtgcgtgtgtgaaccgcttccaaggagccaactgtctcacacacagtgat 448 |
| Db | 318 TGCTGTGACTGGCAGTAGTGATGAACCGCTTTCAGAAAGCAAATTTGTCAGCCACAGTGAT 377 |

| | | | |
|----------|---|--|----------------------|
| QY | 449 | gctgctcgagggaactgctctgcagagatttaccggaagacacaactggtgttttca | 508 |
| | 11 | | |
| Db | 378 | GCCACTCTCGGGGACTGCTCTGCGCAGAGATTTTATGAGAACACCAACTGCGCTTTCAA | 437 |
| QY | 509 | gacatygagctgtgtgtccctcgaggagaccacactcctccctacgaaccacactgtgagta | 568 |
| | 11 | | |
| Db | 438 | GACACTGAGGTGTGTCTCTTGTTGTGAGAACCCCTCCTCTCTTACGAACCGACACTGTGCACAC | 497 |
| QY | 569 | tgtgcgaagtgcagcagacc | 589 |
| | 11 | | |
| Db | 498 | AAGGTCAACTCTGTGAGATC | 518 |
| RESULT | 2 | | |
| ID | V33362 | | |
| AC | V33362; | standard; cDNA to mRNA; 1496 BP. | |
| DT | 02-DEC-1998 | (first entry) | |
| DE | Nucleotide sequence of human beta-OAF065. | | |
| KW | Human; beta-OAF065; stroma cell; antibody; inflammatory; | | |
| KW | cytokine-mediated disease; rheumatism; ulcerative colitis; SS. | | |
| OS | Homo sapiens. | | |
| FM | Key | Location/Qualifiers | |
| FT | CDS | 45..1316 | |
| FT | | /*tag= a | |
| FT | | /product= "human beta-OAF065 protein" | |
| FT | sig_peptide | 45..119 | |
| FT | | /*tag= b | |
| FT | mat_peptide | 120..1313 | |
| FT | | /*tag= c | |
| FT | | /transl_except= (pos:711..713, aa= Pro) | |
| FT | | /transl_except= (pos:714..716, aa= Arg) | |
| PN | W098383304-A1. | | |
| PD | 03-SEP-1998. | | |
| PF | 26-FEB-1998; J00799. | | |
| PR | 27-FEB-1997; JP-043143. | | |
| PA | (ONOY) ONO PHARM CO LTD. | | |
| PI | Fukushima D, Konishi M, Tada H; | | |
| DR | WPI; 98-481205/41. | | |
| DR | P-PSDB: W70387. | | |
| PT | Membrane polypeptide expressed by human stroma cells, and antibodies | | |
| PT | recognising it, for treatment of inflammatory and other | | |
| PT | cytokine-mediated diseases. | | |
| PS | Disclosure; Pages 40-41; 54pp; Japanese. | | |
| CC | This is the nucleotide sequence of the human beta-OAF065, used in | | |
| CC | the method of the invention. The process involves the use of peptides | | |
| CC | expressed by stroma cells, and its antibodies are used for in the | | |
| CC | prevention and treatment of inflammatory and other cytokine-mediated | | |
| CC | diseases such as rheumatism, ulcerative colitis. | | |
| Sequence | 1496 BP; | 388 A; | 360 C; 372 G; 376 T; |

| Query Match | 52.4% | Score 326.6 | DB 1 | Length 1456 |
|-----------------------|----------------|---|----------|-------------|
| Best Local Similarity | 78.2% | Pred. No. 2.4e-86 | | |
| Matches 392 | Conservative 0 | Mismatches 109 | Indels 0 | Gaps 0 |
| QY | 89 | caggataaacacgttcttgtagagacatgycactcaagtcctctactacacagcagc | 148 | |
| | | | | |
| Db | 18 | CAACATATTAATACATTTGATAGAAAGATGGCTTTAAAGTCTCTCTGAAACAAGAA | 77 | |
| QY | 149 | gtgccttcgctgcacatctctctctactccactcggcatgtaagtgaattgagaaac | 208 | |
| | | | | |
| Db | 78 | ACGTTTTCACCTCTTTTAGTATTAATCTAGAGCTATTTTGTCATGTAAAGTGCTGTAA | 137 | |
| QY | 209 | ggaaattgcaggcagcaggaaattcaagaatcgatcttggaaactgltcctctgcacaacg | 268 | |
| | | | | |
| Db | 138 | GGAGCTGTAAACAGCAGAAATTCAGGATCGGTGTGGAATGTGTTCCTCCGCAACCG | 197 | |
| QY | 269 | tcgcgacctgcgcatggaatttgcacaagaattgtgcttcgcgtatacggggaggagatgcacag | 328 | |
| | | | | |
| Db | 198 | TTGTGGCGAGCGATCGAATTGTCTTAAGGAATGTGGCTTCGGCTATACGGAGATGCACAG | 257 | |

| | | | |
|---------|---|--|-----|
| OY | 329 | tgttgcgccctccgagcgacaccggtttcaaggaaagactcgggttttcacagaagttaaacga | 388 |
| Db | 258 | TGTTGTACAGTGCCTGGCGGTGCACAGGTTCCAGAAGAGCATGGGCTTCACGAATATCAAGCCC | 317 |
| OY | 389 | tatgcgacctgtgtcgcttggtgaaccgctttcacagaggccaactgtctacacacagtatg | 448 |
| Db | 318 | TGTTGTGACTGCCAATGGTGTGAACCGCTTTTCAGAAAGGCAATTTGTTCAGCCACAGTGAT | 377 |
| OY | 449 | gcctgtccgggggagcactgcctgcgcagattttacacggaaagaccaactcgtgttgttcaa | 508 |
| Db | 378 | GCCATTGCGCGGGACATGCTTGCACAGATTTATAGGAAGAACAACTTTCGGCTTTCAA | 437 |
| OY | 509 | gacatgagatggttgcctcctgcggagaccacacotcctccctacagacacactgtgatga | 568 |
| Db | 438 | GACATGAGTGTGTGCTTGTGTGGAATACCCTCTCCTCTTACGAAACGCACTGTGCACG | 497 |
| OY | 569 | tatgcacaagtgcagcagcagacc | 589 |
| Db | 498 | AAGGTCAACCTCTGTGAAGATC | 518 |
| RESULT | 3 | | |
| V11422 | | | |
| ID | V11422 standard; cDNA; 371 BP. | | |
| AC | V11422; | | |
| DT | 23-JUL-1998 (first entry) | | |
| DE | Human secreted protein clone AX92_3 cDNA 5'-end. | | |
| KW | Secreted protein; prevention; treatment; gene therapy; ds. | | |
| OS | Homo sapiens. | | |
| PN | WO9801554-A2. | | |
| PD | 15-JAN-1998. | | |
| PF | 07-JUL-1997; U11876. | | |
| PR | 09-JUL-1996; US-677231. | | |
| PA | (GENY) GENETICS INST INC. | | |
| PI | Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM, | | |
| PT | Meiberg D, Racine LA, Spaulding V, Treacy M; | | |
| DR | WPI: 98-110230/10. | | |
| P-SDBS: | M58844. | | |
| FT | Secreted proteins and polynucleotides encoding them - useful to | | |
| PS | prevent, treat and ameliorate medical conditions | | |
| PC | Claim 15; Page 57; 93pp; English. | | |
| CC | V11422-V11424 encode fragments of a novel secreted protein derived from | | |
| CC | clone AX92_3 which was isolated from a human adult testes cDNA library. | | |
| CC | The protein can be used to prevent, treat or ameliorate a medical | | |
| CC | condition, while the polynucleotides can be used for gene therapy. | | |
| QC | Sequence 371 BP; 83 A; 96 C; 107 G; 82 T; | | |

| Query Match | 39.9% | Score 248.8 | DB 1 | Length 371 |
|-----------------------|--|-------------------|---------------|------------|
| Best Local Similarity | 83.8% | Pred. No. 7.8e-64 | | |
| Matches 280 | Conservative | 0 | Mismatches 54 | Indels 0 |
| QY 243 | ctggaactgtgctcctgcgaacaacgtgcygacccgtgcgatgagttgtccaaagaaatgtg | 302 | | |
| Db 1 | CTGGAACCTGTGTTCCTCCGCAACCAAGTGTGGCCACGCATGAGTGTCTAAGGAATGTG | 60 | | |
| QY 303 | gcttcgcatatgaggagatgcacagttgtgtcccttcgaagccgcacccgtgtcaagaag | 362 | | |
| Db 61 | GCTTCGCTATGAGGAGGATGCACAGTGTGTGACGTGGCGGTGCACAGGTCTAAGGAGG | 120 | | |
| QY 363 | actggggtttccagaagtgttaagccaatgtcggactgtgtgcctgtgtgaacgctttcaga | 422 | | |
| Db 121 | ACTGGGCTTCCCAAAAAATGCAAGCCCTCTCTGACACTCGCAGTGGTGAACCCCTTTACA | 180 | | |
| QY 423 | ggggcaactgttcacaaccacagtgatgtctgtcggggagactgtcctgcacagatttacc | 482 | | |
| Db 181 | AGGCCAATTGTTCACCCACCAAGTAGTGCATCTGGGGGAGACTGCTTGCACANAATTTATA | 240 | | |
| QY 483 | ggaagacccaactgttgtttgtttcaagaacaggaatgtgtgcccctgcggagaaacccc | 542 | | |
| Db 241 | GGAAGACCAAAACTTGTGGGCTTTTCAAAACATGGAAGTGTGTGCTTTGGAAACCTCTCTC | 300 | | |
| QY 543 | ctccctacgaaccacactgtgagtatgtgccaa | 576 | | |

```
Db 301 CTCCTTAGCAGCCGCACTGTGCCAGCAAGTCGA 334
|||||
RESULT 4
V11423
ID V11423 standard; cDNA; 181 BP.
AC V11423.
DT 23-JUL-1998 (first entry)
DE Human secreted protein clone AX92_3 cDNA internal fragment.
KW Secreted protein; prevention; treatment; gene therapy; ds.
OS Homo sapiens.
PN M09801554-A2.
PD 15-JAN-1998.
PF 07-JUL-1997; 011876.
PR 09-JUL-1996; US-677231.
PA (GENY ) GENETICS INST INC.
PI Bowman M, Evans C, Jacobs K, Lavallie ER, McCoy JM,
PR Merberg D, Racie LA, Spaulding V, Treacy M;
DR WPI: 98-110230/10.
P-PSDB: W58844.
PT Secreted proteins and polynucleotides encoding them - useful to
PS prevent, treat and ameliorate medical conditions
PS Claim 13: Page 57; 93pp. English.
CC V11422-V11424 encode fragments of a novel secreted protein derived from
CC clone AX92_3 which was isolated from a human fetal brain cDNA library.
CC The protein can be used to prevent, treat or ameliorate a medical
CC condition, while the polynucleotides can be used for gene therapy.
SQ Sequence 181 BP; 43 A; 47 C; 34 G; 57 T;

Query Match 11.3%; Score 70.2; DB 1; Length 181;
Best Local Similarity 85.7%; Pred. No. 1e-11;
Matches 78; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

OY 474 gatttacggagagccaactggtgttccaagacatgagtgccctggag 533
|||
DB 88 GATTTATAGGAGACGAAGACTGTGCTTCAGACATGATGTGTGGAG 147
|||||
OY 534 accacactcctccactcagacacacactgtga 564
|||
DB 148 ACCCTGCTCTCCTTACGACCGCACTCTCA 178
|||||

RESULT 5
O10388/c
ID O10388 standard; cDNA; 1912 BP.
AC O10388.
DT 11-APR-1991 (first entry)
DE Encodes skin-specific murine Retinoic Acid Receptor-gamma.
KW mouse; retinoic acid receptor; mRAR-gamma; embryogenesis; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT cds 316..1892
FT /tag= a
FT /product= murine RAR-gamma

EP-411323-A.
PD 06-FEB-1991.
PF 29-JUN-1990; 112469.
PR 30-JUN-1989; US-374690.
PA (INRM ) INSERM INST NAT SANTE.
PA (SOUT ) LES LABS SQUIBB SA.
PI Chambon P, Kastner P, Krust A, Petkovich M, Zelent A;
PI Leroy P, Mendelsohn C, Staub A;
DR WPI: 91-038271/06.
DR P-PSDB: R105349.
PT Novel human and mouse retinoic acid receptors - encode proteins
PT used to assay for agonists and antagonists
PS Disclosure: Fig 1; 33pp. English.
CC An 11.5 day-old total mouse embryo lambda gt10 cDNA library was
CC screened with labelled human RAR-alpha and -beta probes. Positive
CC clones were isolated, mapped and sequenced. A set of clones was
```

```
CC identified having a lesser degree of homology with either RAR-alpha
CC or -beta than murine RAR-alpha or mRAR-beta clones, respectively.
CC The cDNA-deduced amino acid sequence of this third set of clones
CC was obviously related to mRAR-alpha and mRAR-beta. This new member
CC of the mouse RAR subfamily was designated RAR-gamma. The
CC mRAR-gamma cDNA clone was used to clone 7 different murine
CC RAR-gamma cDNA isoforms, generated by alternative splicing of at
CC least 7 exons. The isoforms were found to differ in their 5'-UTR
CC sequences and in their N-terminal A region, which is known to be
CC important for differential trans-activation by other nuclear
CC receptors.
CC See also R10547-8, Q10389 and Q10405-8.
SQ Sequence 1912 BP; 427 A; 597 C; 534 G; 354 T;

Query Match 5.6%; Score 35; DB 1; Length 1912;
Best Local Similarity 47.5%; Pred. No. 0.53;
Matches 104; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

OY 207 ccgagatgcagcagcagaatccaagatcgaatctggaactgtctctcgaac 266
|||
DB 233 CCTGGCCTGGAGAGGCTCCGTCGCCACCTGCTGCTGCCAGTGGGTTCCAGAGATC 174
|||||
OY 267 agtgcggaactgcatgagatgttccaagaaatgtgcttgctatggagagatgcac 326
|||
DB 173 CCTGGCCTCTGCAGAGAGTCCCTCAATACATGCGCCAGCGTAGGTGAGACCTG 114
|||||
OY 327 agtgtgtgcctgcagcagccgacacggttcaaggaagactggttccagaagtgaagc 386
|||
DB 113 GTGGAACCTGCTGGAGGGAGGGCCGGGGAGAGAGGCCCGCAGACGCCGAGGTCCCGGC 54
|||||
OY 387 catgtcggactgtgcgtgtgaacgcgttccagaagg 425
|||
DB 53 GTCGGGCGGTCTCGGGGATGAGACCGCGGTGGGGGG 15
|||||

RESULT 6
T05202/c
ID T05202 standard; cDNA; 1913 BP.
AC T05202.
DT 13-JUN-1996 (first entry)
DE Murine retinoic acid receptor gamma cDNA.
KW Retinoic acid receptor; mouse; RAR; RAR-gamma; transcription factor;
KW skin; ss.
OS Mus musculus.
FH Key Location/Qualifiers
FT cds 316..1692
FT /tag= a
FT /product= retinoic acid receptor gamma protein
FT misc_feature 316..498
FT /tag= b
FT /note= "hRAR-gamma A region"
FT misc_feature 499..582
FT /tag= c
FT /note= "hRAR-gamma B region"
FT misc_feature 583..780
FT /tag= d
FT /note= "hRAR-gamma C region"
FT misc_feature 781..918
FT /tag= e
FT /note= "hRAR-gamma D region"
FT misc_feature 919..1578
FT /tag= f
FT /note= "hRAR-gamma E region"
FT misc_feature 1579..1689
FT /tag= g
FT /note= "hRAR-gamma F region"
FT EP-683227-A1.
PD 22-NOV-1995.
PF 29-JUN-1990; 110924.
PR 30-JUN-1989; US-374690.
PA (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
```

PA (BRIM) INST RECH SOUTIB.
 PA (BRIM) LES LAB SOUTIB SA.
 PI Chamoun P, Kastner P, Krust A, Leroy P, Mendelsohn C;
 PI Petkovich M, Stauda, Zelent A;
 DR MPI: 95-394349/51.
 DR P-PSDB: R84726.
 PT New cDNA encoding the human retinoic acid receptor gamma - useful
 PT for identifying agonists and antagonists of RAR-gamma and detecting
 PT expression in normal and diseased tissue
 PS Example 1: Fig 1: 34pp: English
 CC This sequence represents the cDNA encoding the mouse retinoic acid
 CC receptor gamma (RAR-gamma). This sequence was used in the isolation of
 CC the human RAR-gamma sequence. RAR-gamma is a retinoic acid inducible
 CC transcription factor, and is expressed specifically in the skin where it
 CC is the predominant RAR. The C and E regions of the protein are the most
 CC highly conserved regions through the RAR's of a species. The C region is
 CC the DNA binding domain, and the E region is the ligand binding domain.
 CC The B region is also well conserved within species. The N and C terminal
 CC regions of segment D are conserved, whereas the central segment of this
 CC region is not. The D region is thought to act as a hinge region.
 CC Regions A and F are not conserved within a given species. The A, B, D
 CC and F regions are conserved across species, suggesting that they have
 CC specific functions, but are not required for the RAR's specific
 CC physiological roles. The protein is used to detect competitive and
 CC non-competitive agonists and antagonists. This sequence can be used to
 CC detect receptor expression in normal and diseased tissues and cultures.
 SQ Sequence 1913 BP: 427 A; 600 C; 531 G; 355 T;

Query Match 5.6%; Score 35; DB 1; Length 1913;
 Best Local Similarity 47.5%; Pred. No. 0.53;
 Matches 104; Conservative 0; Mismatches 115; Indels 0; Gaps 0;

QY 207 ccggaagatcgcagcagcaggaattcaagatgatctgaactgtctcttgcacaac 266
 DB 233 CCTGGCCTGGAGAGCTCCGCCACCCCTCGTCTGTCGAGTGTCCAGAGATC 174
 QY 267 agtgcgacacctgcagatgtgtccaaagatgtgcttcgagtgaggagatgcac 326
 DB 173 CCGTGTGGCCTTCGACAGAGATCCCTCTCATATCTGAGCCAGCTAGAGGAGACTG 114
 QY 327 agtgtgtccctcagcgcgcacacggttcaagagaagatcgtgggttccagaagtgtgaagc 386
 DB 113 GTGGAACCTCTGTGAGGAGGAGGCGGAGGAAGAGCCGACAGCCCGAGGTCCCGGC 54
 QY 387 catgtgcgacacctgtgctgtgaacaccttcaagagg 425
 DB 53 GTGCGCGGTCTCGGGGATGACACCGCGGTGGGGGG 15

RESULT 7
 V39359 standard; cDNA; 1013 BP.
 AC V39359;
 DT 21-SEP-1998 (first entry)
 DE Humanised anti-HM1.24 antibody encoding cDNA.
 KW Mouse; human; humanised; anti-HM1.24 antibody; myeloma; FR; CDR;
 KW framework region; complementarity determining region; antigenicity; ss.
 OS Synthetic.
 OS Mus sp.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT CDS 23..565
 FT /tag= a
 PN MO9814560-A1.
 PD 09-APR-1998.
 PF 03-OCT-1997; J03553.
 PR 04-OCT-1996; JP-264756.
 PA (CHUS) CHUGAI SEIYAKU KK.
 PI Koishihara Y, Kosaka M, Ohtomo T, Ono K, Tsuchiya M,
 PI Yoshimura Y;
 DR MPI: 98-286421/25.
 DR P-PSDB; W62207.

PT Humanised anti-HM1.24 antibody - for treatment of myeloma
 PS Claim 81: Page 151-153: 210pp: Japanese.
 CC A humanised anti-HM1.24 antibody has been developed which comprises
 CC human L and H chain C regions, and L and/or H chain V regions
 CC containing material originating in mouse anti-HM1.24 antibody. The V
 CC regions contain framework (FR) regions of human origin and
 CC complementarity determining regions (CDR) of mouse origin, leading to
 CC a reshaped humanised antibody. The C regions are human Ck (L-chain) and
 CC human C gamma (especially C gamma 1) (H-chain). The FR regions of the
 CC L chain V region are derived from human subtype HsG1 (e.g. from human
 CC antibody RE1) and the FR regions of the H chain V region are derived
 CC from human subtype HsG1 (e.g. FR1-3 from human antibody Hg3 and FR4
 CC from human antibody JH6). The present sequence encodes an antibody
 CC polypeptide from the present invention. The antibodies are used for
 CC the treatment of myeloma, especially by injection, intravenously,
 CC intramuscularly or subcutaneously. The antibodies are used at 0.01-1000
 CC (especially 5-100) mg/kg body weight. The humanised antibody has low
 CC antigenicity and is therefore effective therapeutically in humans.
 SQ Sequence 1013 BP: 232 A; 254 C; 314 G; 213 T;

Query Match 5.5%; Score 34.2; DB 1; Length 1013;
 Best Local Similarity 83.0%; Pred. No. 0.71;
 Matches 39; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 576 agtgcgacgagacgttaaaaaaaaaaagaaacaaacaaaaa 622
 DB 953 AGGAGAGCGACACTTAAAAAANAAAAAANAAAAAANAAAAA 999

RESULT 8
 V07579 standard; cDNA; 1013 BP.
 AC V07579;
 DT 20-NOV-1998 (first entry)
 DE Protein recognised by anti-human HM1.24 antibody nucleotide sequence.
 KW ss; cytotoxic antibody; anti-human HM1.24; lymphocytic tumours;
 KW pre-B lymphoma; Burkitt's lymphoma; T-cell tumour; B-cell tumour.
 OS Homo sapiens
 FH Key Location/Qualifiers
 FT CDS 23..565
 FT /tag= a
 FT /product= "Cytotoxic protein"
 PN MO9835698-A1.
 PD 20-AUG-1998.
 PF 12-FEB-1996; J00568.
 PR 12-FEB-1997; JP-041410.
 PA (CHUS) CHUGAI SEIYAKU KK.
 PI Koishihara Y, Yoshimura Y;
 DR MPI: 98-456869/39.
 DR P-PSDB; W65771.

PT Treatment of lymphocytic tumours using cytotoxic antibody - binding
 PT to specific antigen such as HM1.24 and effective against T-cell
 PT tumours and B-cell tumours other than myeloma
 PS Claim 1: Page 44-45: 82pp: Japanese.
 CC The protein having the nucleotide sequence below is bound specifically
 CC by a cytotoxic antibody which can be used in the treatment of lymphocytic
 CC tumours, including T-cell tumours and B-cell tumours other than myeloma.
 CC The antibody is preferably monoclonal and has ADCC or CDC type
 CC cytotoxicity. It may be chimeric or humanised, and preferably contains
 CC a human antibody constant region C gamma (such as C gamma 1 or C
 CC gamma 3). A preferred antibody is an anti-human HM1.24 antibody or an
 CC antibody which binds to an epitope recognising anti-human HM1.24
 CC antibody. The cytotoxic antibody is useful in the treatment of
 CC lymphocytic tumours such as acute or chronic B lymphocytic leukemia,
 CC pre-B lymphoma, Burkitt's lymphoma, or acute or chronic T lymphocytic
 CC leukaemia.
 SQ Sequence 1013 BP: 232 A; 257 C; 311 G; 213 T;

Query Match 5.5%; Score 34.2; DB 1; Length 1013;
 Best Local Similarity 83.0%; Pred. No. 0.71;
 Matches 39; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 576 agtgcagcagcagccttaaaaaaagaaaaaacaacaaaaa 622
DB 953 AGGAGAGCAGCACCCTTAAAAAAGAAAAAAGAAAAA 999

RESULT 9

VS9067
ID V59067 standard: cDNA; 2634 BP.

AC V59067;

DT 02-FEB-1999 (first entry)

DE Bax inhibitor Bi-1 cDNA

KM Bax inhibitor: Bi-1; human; apoptosis; ss.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 73..786

FT CDS /tag= a

PN MO9840397-A1.

PD 17-SEP-1998.

PF 13-MAR-1998; 005015.

PR 14-MAR-1997; 05-818514.

PA (BURN-) BURNHAM INST.

PI Reed JC, Xu Q;

DR WPI: 98-531519/45.

P-PSDB: W73136.

PT Bax inhibitor proteins, Bi-1 and Bi-2 - useful e.g. to modulate

PT cellular apoptotic activity or identify agents altering Bi-1 or Bi-2

PT binding which can modulate apoptotic activity

PS Claim 2: Page 61-63; 80pp; English.

CC This cDNA clone codes for an inhibitor protein, termed Bi-1 (see

CC W73136), of the pro-apoptotic protein Bax. Nucleic acids encoding

CC Bi-1 and Bi-2 (see VS9068) were identified by suppression of

CC Bax-induced death of yeast cells transformed to express human Bax.

CC A human HepG2 cDNA library was used for library screening. The

CC invention provides vectors, optionally expression or viral vectors,

CC containing Bi nucleic acids, and host cells containing these

CC vectors. The nucleic acids encoding Bi-1/Bi-2 can be used to

CC increase expression of these proteins in cells, or antisense

CC molecules prepared from them used to decrease expression. In

CC these ways, cellular apoptotic activity may be modulated (claimed).

CC The nucleic acids and complementary sequences are also useful as

CC probes to detect Bi-encoding nucleic acid molecules in samples.

CC Sequence 2634 BP: 632 A; 626 C; 597 G; 779 T;

SO

Query Match

Best Local Similarity 5.4%; Score 33.8; DB 1; Length 2634;

Matches 41; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

QY 570 gtgcgaagtgagcagccttaaaaaaagaaaaaacaacaaaaa 622

DB 2579 GTGCCAAGATGACAGATATTAATAAAAAAAAAAAAAA 2631

RESULT 10

VS9114

ID V59114 standard: cDNA; 1016 BP.

AC V59114;

DT 14-DEC-1998 (first entry)

DE Nucleotide sequence of protein bound by Anti-HM1.24 antibody.

KM ss; Anti-HM1.24; antibody; lymphocyte activation inhibitor;

KW autoimmune disease; organ transplant; allergy.

OS Homo sapiens.

FH Key Location/Qualifiers

FT CDS 23..565

FT CDS /tag= a

PN MO9837913-A1.

PD 03-SEP-1998.

PF 27-FEB-1998; J00831.

PR 28-FEB-1997; JP-045663.

PA (CHUS) CHUGAI SEIYAKU KK.

PI Koshinara Y;

DR WPI: 98-480937/41.

DR P-PSDB: W77292.

PT Lymphocyte activation inhibitor comprises antibodies, particularly

PT anti-HM1.24 antibody - for preventing and treating autoimmune

PT diseases, rejection reactions in organ transplant or allergy

PS Claim 1: Page 38-39; 53pp; Japanese.

CC The Anti-HM1.24 antibody can be used in the production of lymphocyte

CC activation inhibitors. These inhibitors can be used for the prevention

CC and treatment of autoimmune diseases, rejection reactions in organ

CC transplant or allergy. Administration is non-oral, e.g. by intra

CC venous and intramuscular injection, local or systemic.

SO Sequence 1016 BP: 233 A; 258 C; 312 G; 213 T;

Query Match

Best Local Similarity 5.4%; Score 33.4; DB 1; Length 1016;

Matches 37; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 580 gcagcagcagccttaaaaaaagaaaaaacaacaaaaa 622

DB 960 GCACACACACCTTAAAAAAGAAAAAAGAAAAA 1002

RESULT 11

T72107

ID T72107 standard: cDNA to mRNA; 3692 BP.

AC T72107;

DT 13-FEB-1998 (first entry)

DE Rat semaphorin 2 gene.

KM Semaphorin 2; central nerve extension; rat; human; inhibitor;

KW central nerve regeneration promoter; ss.

OS Rattus norvegicus.

FH Key Location/Qualifiers

FT 5'UTR 1..18

FT CDS /tag= a

FT CDS /tag= b

FT 3'UTR 2683..3653

FT polyA_signal 3654..3692

FT /tag= d

PN MO9720928-A1.

PD 12-JUN-1997.

PF 02-DEC-1996; J03517.

PR 31-OCT-1996; JP-307205.

PR 06-DEC-1995; JP-345187.

PA (SUMO) SUMITOMO PHARM CO LTD.

PI Kikuchi K, Kimura T;

DR WPI: 97-319775/29.

P-PSDB: W19856.

PT Semaphorin 2 and the gene encoding it - also inhibitors of its

PT action which can be used as promoters of central nerve regeneration

PS Claim 1: Page 61-63; 10pp; Japanese.

CC This sequence represents the rat semaphorin 2 gene. The protein encoded

CC by this sequence, and the human semaphorin 2 protein (see W19857) can be

CC used for screening possible candidates for activity as semaphorin 2

CC inhibitors. Proteins containing all or part of the semaphorin 2 sequence

CC are useful as central nerve extension inhibitors. Semaphorin 2 inhibitors

CC identified by screening can be used as promoters of central nerve

CC regeneration.

SO Sequence 3692 BP: 689 A; 1126 C; 1122 G; 755 T;

Query Match

Best Local Similarity 5.3%; Score 33.2; DB 1; Length 3692;

Matches 59; Conservative 0; Mismatches 43; Indels 0; Gaps 0;

QY 521 gtgcctgcagagcagcagccttcctcagcaacacactgtgagtgcaagtg 580

DB 3582 GTGGCAGCGATGATCCAGAGCCTCTCTCAGCCCTATCCCTGTCAATAAAGCTTGT 3641

QY 581 cagcagaccccttaaaaaaagaaaaaacaacaaaaa 622

DB 3582 GTGGCAGCGATGATCCAGAGCCTCTCTCAGCCCTATCCCTGTCAATAAAGCTTGT 3641

DB 3642 CATCCACGCGCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3683

RESULT 12

ID T89343 standard; cDNA; 2083 BP.

AC T89343.

DT 11-MAR-1998 (first entry)

DE Human p62 cDNA #1.

KW p62; cytoplasmic; T cell; B cell; development; activation; modulation; cellular response; cell proliferation; autoimmune disease; ss.

OS Homo sapiens.

PH Key Location/Qualifiers

FT CDS 67..1389

FT /tag= a

FT /product= p62

PN MO9722255-A1.

PD 26-JUN-1997.

PR 11-DEC-1996; U19944.

PS 19-DEC-1995; US-574959.

PA (DAND) DANA FARMER CANCER INST INC.

PI Young I, Shin J, Strominger JL, Vadmudi RK;

DR WPI: 97-341351/31.

PT P-PSDB: W31182.

CC cDNA encoding p62 and p160 and corresponding proteins - used in the treatment of autoimmune disease and for T and B cell proliferation, e.g. for treatment of tumours

PS Claim 4: Fig 1: 175pp; English.

CC This cDNA sequence encodes a novel p62 cytoplasmic polypeptide which is expressed in a variety of tissues including heart, brain, placenta, lung, liver, skeletal muscle, kidney and pancreas. This polypeptide is capable of modulating T or B cell development and/or T or B cell activation e.g. by modulation of cell activity. It is also capable of modulating degradation of cellular proteins e.g. cell cycle regulatory proteins stimulating expression of cell cycle dependent kinase inhibitors and arresting cell cycle progression at specific boundaries to thereby modulate cell proliferation. As p62 acts to boost B cell response and may be used to treat disorders where this is beneficial, e.g. infections by pathogenic microorganisms, e.g. bacteria, viruses and protozoans. p62 can be used to expand T cell populations for treating infectious diseases or cancer, e.g. the resulting cells may be transduced to render them resistant to HIV infection. Inhibitors of p62 can be used to reduce B or T cell responses and may be used to treat a variety of autoimmune diseases, e.g. diabetes mellitus, arthritis, multiple sclerosis allergic reactions, Crohn's diseases etc.

CC Sequence 2083 BP; 447 A; 588 C; 608 G; 440 T;

SO

Query Match 5.3%; Score 33; DB 1; Length 2083;

Best Local Similarity 79.6%; Pred. No. 2.1;

Matches 39; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

OY 575 aagtygcagcagaccttaaaaaaagaacaaacacaaac 623

DB 2027 AATGTGCACTACGACTTAAACCAATMAAAAAAAAAAAAAAAAAAAC 2075

RESULT 13

ID T73502 standard; cDNA to mRNA; 1868 BP.

AC T73502.

DT 16-SEP-1997 (revised)

DT 26-AUG-1997 (first entry)

DE Poplar 1-aminocyclopropane-1-carboxylic acid synthase gene PNACC52.

KW Ozone; induction; exposure; resistance; transgenic plant; ACC;

KW 1-aminocyclopropane-1-carboxylic acid biosynthesis; ds.

OS Populus nigra.

PH Key Location/Qualifiers

FT CDS 119..1579

FT /tag= a

FT /product= PNACC52

PN J09075088-A.

PD 25-MAR-1997.

PF 07-SEP-1995; 254510.

PR 07-SEP-1995; JP-254510.

PA (TOYT) TOYOTA JIDOSHA KK.

DR WPI: 97-239270/22.

DR P-PSDB: W21755.

PT Ozone-induced 1-amino-cyclopropane-1-carboxylic acid synthetic enzyme genes - from poplar tree, are useful for generating ozone-resistant trees and pollution clean-up trees

PT enzyme genes

PS Claim 2: Pages 9-11; 12pp; Japanese.

CC This sequence, designated PNACC52, is a 1-aminocyclopropane-1-carboxylic acid (ACC) synthase gene isolated from poplar trees which had been exposed to ozone. Exposure conditions were: 0.1 ppm ozone for 6 hours after one hour acclimation at 25 degrees C, 70 % relative humidity, 30 k lux and air current 30 cm/second. This gene will be useful for breeding air pollutant ozone-resistant trees, especially poplar.

CC (Note: Record has been revised in order to correct the Organism Source field).

CC Sequence 1868 BP; 590 A; 366 C; 405 G; 507 T;

SO

Query Match 5.3%; Score 32.8; DB 1; Length 1868;

Best Local Similarity 67.6%; Pred. No. 2.3;

Matches 46; Conservative 0; Mismatches 22; Indels 0; Gaps 0;

OY 555 cacactgagtgatgtgtccagatgcagcagaccttaaaaaaagaacaaac 614

DB 1774 CAATTGAAATTCAGATGATATTGTCTCTCTTTAAAAAAAAAAAAAAAA 1833

OY 615 aacaaaaa 622

DB 1834 AAAAAAAAA 1841

RESULT 14

ID 054832 standard; DNA; 1227 BP.

AC 054832.

DT 19-JUL-1994 (first entry)

DE Sequence of clone W225 encoding mouse gonadotropin-releasing hormone receptor (GnRH-R).

DE Gonadotropin-releasing hormone receptor; GnRH; diagnosis;

KW contraceptive; reproductive disorders; ss.

OS Mus musculus.

PH Key Location/Qualifiers

FT CDS 43..1023

FT /tag= a

PN M09400590-A.

PD 06-JAN-1994.

PF 22-JUN-1993; U05965.

PR 23-JUN-1992; US-904072.

PR 21-JUN-1993; US-080386.

PA (UYN) UNIV NEW YORK MT SINAI SCHOOL MEDICINE.

PI Sealton SC;

DR WPI: 94-026225/03.

DR P-PSDB: R47164.

PT Gonadotropin-releasing hormone receptor genes and proteins - for expression of GnRH and screening and identifying GnRH (ant)agonists, for diagnosis and therapy of reproductive disorders and for contraception

PS Claim 8: Page 44-46; 73pp; English.

CC The longest ORF of 054832 encodes a 327 AA protein of about 37,000 MW (R47164). The product is murine GnRH-R. Three consensus N-linked glycosylation sites are present, two in the N-terminus and one in the first extracellular loop. Hydrophobicity analysis of the deduced protein reveals seven stretches of highly hydrophobic AAs with 20-30% sequence similarity to other G-protein receptors, with the highest degree of homology to the interleukin-8 receptor.

CC Sequence 1227 BP; 337 A; 308 C; 246 G; 336 T;

SO

Query Match 5.2%; Score 32.6; DB 1; Length 1227;

Best Local Similarity 80.9%; Pred. No. 2.2;

This page blank (uspio)

| Db | 236 | 1GCAAAAGTGGGACACTGGCATGGAGTTGTCCAAAGAAATGTGGCTTGGCCATATGGGGAG | 295 |
|-----------------------|--|--|--|
| Oy | 320 | gatgcacagttgtgtcctctgcagccgcgcagcgggttcagaagactgggtttccagaag | 379 |
| Db | 296 | GATGCACAGTGTGTGCTCCCTGCAGGCCGCACCGTTTCAAGAAAGACTGGGCTTCCAGAAG | 355 |
| Oy | 380 | tgt 382 | |
| Db | 356 | TGT 358 | |
| RESULT | 3 | | |
| LOCUS | AA575933/c | | |
| DEFINITION | AA575933 | 520 bp | mRNA EST 09-SEP-1997 |
| ACCESSION | nm06c10.s1 | NCI_CGAP_Br3 | Homo sapiens cDNA clone IMAGE:1072242 |
| NID | AA575933 | | similar to SW:COX1_HUMAN P00395 CYTOCHROME C OXIDASE POLYPEPTIDE I |
| KEYWORDS | AA575933 | | ; mRNA sequence. |
| SOURCE | G2350448 | | EST. |
| ORGANISM | human. | | human. |
| REFERENCE | Homo sapiens | | |
| AUTHORS | Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; | | |
| TITLE | Primates; Catarrhini; Homnidae; Homo. | | |
| JOURNAL | NCI-CGAP | http://www.ncbi.nlm.nih.gov/ncicgap. | |
| COMMENT | 1 (bases 1 to 520) | | |
| | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), | | |
| | Tumor Gene Index | | |
| | Unpublished (1997) | | |
| | Contact: Robert Strausberg, Ph.D. | | |
| | Tel: (301) 496-1550 | | |
| | Email: Robert_Strausberg@nih.gov | | |
| | Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R. | | |
| | Emmert-Buck, M.D., Ph.D. | | |
| | CDNA Library Preparation: Stratagene, Inc. | | |
| | CDNA Library Arrayed by: Greg Lennon, Ph.D. | | |
| | DNA Sequencing by: Washington University Genome Sequencing Center | | |
| | Clone distribution: NCI-CGAP clone distribution information can be | | |
| | found through the I.M.A.G.E. Consortium/ILNL at: | | |
| | www-bio.llnl.gov/bbrp/image/image.html | | |
| FEATURES | Insert Length: 715 | Std Error: 0.00 | |
| source | Seq primer: -40m13 fwd. ET from Amersham | | |
| | High quality sequence stop: 322. | | |
| | Location/Qualifiers | | |
| | 1..520 | | |
| | /organism="Homo sapiens" | | |
| | /note="Organ: breast; Vector: Bluescript SK-; Site.1: | | |
| | EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer: | | |
| | 0150 dt. Ductal breast tumor. 5' adaptor sequence: 5' | | |
| | GATTCGCGACGACGAG 3' 3' adaptor sequence: 5' | | |
| | GTTCAGTGTGTGTGTGTGTGTGT 3' Average insert size: 0.9 kb." | | |
| | /db_xref="taxon:9606" | | |
| | /clone IMAGE:1072242" | | |
| | /clone_lib="NCI_CGAP_Br3" | | |
| | /sex="female" | | |
| | /tissue_type="breast tumor" | | |
| | /lab_host="SOLR (kanamycin resistant)" | | |
| | <1..>520 | | |
| BASE COUNT | 139 a | 86 c | 133 g 162 t |
| ORIGIN | | | |
| Query Match | 7.3% | Score 45.6: | DB 22; Length 520; |
| Best Local Similarity | 52.0% | Pred. No. 0.0095: | |
| Matches 102; | Conservative 0; | Mismatches 94; | Indels 0; Gaps 0; |
| Oy | 425 | gcccaactgtcatcacaccacgatgatgtctgtcgtcgggagctgcgtccagatattaccgg | 484 |
| Db | 196 | GCTGACTCCGCCACACTCCACGGAGGAATATGTAATAATGATCTCTGCGAGTCTTGAGCCC | 137 |
| Oy | 485 | aagaccaaactgtgtgttttcaagacatgatgagttgtgtcctctgcggagaccactctct | 544 |

[illegible]

LOCUS A117203 417 bp mRNA EST 17-FEB-1997
DEFINITION mt13604.r1 Soares mouse 3nbms Mus musculus cDNA clone 620558 5',
ACCESSION A117203
NID 91759651
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; Eukaryotes; Chordata; Vertebrata; Mammalia; Muridae; Murinae;
Mus.
REFERENCE 1 (bases 1 to 417)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Geisels,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maria M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:381782
Seq primer: -28M13 rev2 from Amersham
High quality sequence stop: 384.
Location/Qualifiers
1..417
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Vector: p7713D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTCACCAATCGAATGAGAGGCGCCGCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p7713 vector. RNA
provided by Dr. Bertrand Jordan. Library went through
three rounds of normalization, and was constructed by
Bento Soares and M.Fatima Bernaldo."
/DB_xref="taxon:10090"
/clone_lib="Soares mouse 3nbms"
/clone_id="620958"
/sex="male"
/tissue_type="Spleen"
/dev_stage="4 weeks"
/lab_host="DH10B"
1..>417
BASE COUNT 129 a 107 c 98 g 83 t
ORIGIN
mRNA
Query Match 6.3%; Score 39.2; DB 16; Length 417;
Best Local Similarity 69.7%; Pred. No. 0.51;
Matches 53; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

RESULT 6
R53741

LOCUS R53741 421 bp mRNA EST 18-MAY-1995
DEFINITION Y102H05.r1 Homo sapiens cDNA clone 138105 5',
ACCESSION R53741
NID 9815643
KEYWORDS EST.
SOURCE human clone-138105 library-Soares placenta Nb2HP vector-p7713D
(Pharmacia) with a modified polylinker host-DH10B (ampicillin
resistant) primer-M13Rpl Rsite1-Not I Rsite2-Eco RI female placenta
obtained at birth (full term). 1st strand cDNA was primed with a
Not I - oligo(dT) primer [5'
AACGGAAGAATTCGCGCCGACGAGAAATTTTCTTTTCTTTTCTTTTCTTTTCTTTT
3'], double-stranded
cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not
I and cloned into the Not I and Eco RI sites of the modified p7713
vector. Library went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bernaldo.
Homo sapiens
Eukaryotae; Metazoa; Eumetazoa; Bilateria; Coelomata;
Deuterostomia; Chordata; Vertebrata; Gnathostomata; Osteichthyes;
Sarcopterygii; Chonata; Tetrapoda; Amniota; Mammalia; Theria;
Eutheria; Archonta; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 421)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaskis,E., Waterston,R., Williamson,A., Wohlmann,P. and
Wilson,R.
The WashU-Merck EST Project
JOURNAL Unpublished (1995)
COMMENT Contact: Wilson R.
WashU-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
High quality sequence stops: 321
Source: IMAGE Consortium, LNL
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Location/Qualifiers
1..421
/organism="Homo sapiens"
/clone_id="138105"
BASE COUNT 140 a 113 c 87 g 76 t 5 others
ORIGIN
Query Match 6.2%; Score 38.8; DB 10; Length 421;
Best Local Similarity 61.6%; Pred. No. 0.66; Mismatches 38; Indels 0; Gaps 0;
Matches 61; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

RESULT 7
AA789781

LOCUS AA789781 306 bp mRNA EST 06-FEB-1998
DEFINITION vt79c07.r1 Barziled mouse irradiated colon MRLR5 Mus musculus cDNA
clone 1177356 5' similar to gb:X69181 605 RIBOSOMAL PROTEIN L31
(HUMAN); mRNA sequence.
ACCESSION AA789781
NID 92849901
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryotae; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;

[illegible]

| | |
|-----------------------|---|
| AUTHORS | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap . |
| TITLE | National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index |
| JOURNAL | Unpublished (1997) |
| COMMENT | Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_Strausberg@nih.gov Unknown library type Seq primer: -40bp from Gibco. |
| FEATURES | Location/Qualifiers |
| SOURCE | 1..255 /organism="Homo sapiens" /note="Organ: lymph node; Vector: pCMV-Sport6; Site_1: Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.25 kb. Life technologies catalog #: 11547-015" /db_xref="taxon:9606" /clone_image="2000296" /clone_lib="NCI-CGAP_Lym12" /tissue_type="lymphoma, follicular mixed small and large cell" /lab_host="DH10B" |
| BASE COUNT | 43 a 60 c 54 g 98 t |
| ORIGIN | /lab_host="DH10B" |
| Query Match | 6.0%; Score 37.6; DB 31; Length 255; |
| Best Local Similarity | 57.8%; Pred. No. 1.3; |
| Matches | Conservative 0; Mismatches 49; Indels 0; Gaps 0; |
| Oy | 506 caagacatgtagtgctgtgcctcgaggagaccacctcctccctacgaacacactgtag 565 Db 116 CAAAAGCGGGAGTGTCAGTGTGCAGCCGAATGCCCCCCCTGCCCTCCACCTGGGTGACAG 57 Oy 566 tgatgtgcagaatggcagcacgacctttaaaaaaaagaaaaaaaacaacaaaca 621 Db 56 AATGAGACCCTGCTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1 |
| RESULT | 9 |
| LOCUS | C62130 360 bp mRNA EST 29-AUG-1997 |
| DEFINITION | C.elegans cDNA clone yk270b5 : 5' end, single read, mRNA sequence. |
| ACCESSION | C62130 |
| NID | 92420835 |
| KEYWORDS | EST; EST(expressed sequence tag); \ |
| SOURCE | Caenorhabditis elegans (Strain:CBL489 him-8(e1489)) varied hermaphrodite, male whole animal cDNA to mRNA, clone_lib:yuj1 Kohara unpublished cDNA library clone:yk270b5. |
| ORGANISM | Caenorhabditis elegans Eukaryota; Metazoa; Nematozoa; Secernentea; Rhabdilitia; Rhabdilitida; Rhabdilitae; Rhabditoidea; Rhabdilitidae; Peloderinae; Caenorhabditis. |
| REFERENCE | 1 (sites) Kohara,Y., Notohashi,T., Tabara,H., Shin-i.T., Watanabe,H., Sano,M., Miyata,A., Ohba,T., Mitani,Y., Uesugi,H., Sugitara,I., Obara,M., Sugimoto,A., Iida,K. and Nishigaki,A. Expression map of the C.elegans genome |
| AUTHORS | Unpublished (1987) |
| TITLE | 2 (bases 1 to 360) |
| REFERENCE | Kohara,Y. |
| JOURNAL | Direct Submission Submitted (22-JUL-1997) to the DDBJ/EMBL/Genbank databases. Yuj1 Kohara, National Institute of Genetics, Gene Network Lab: Yata 1111, Mishima, Shizuoka 411, Japan (E-mail:ykohara@lab.nig.ac.jp, Tel:+81-559-81-6854, Fax:-81-559-81-6855) |
| FEATURES | Location/Qualifiers |
| SOURCE | 1..360 /organism="Caenorhabditis elegans" /strain="CBL489 him-8(e1489)" /db_xref="taxon:6239" /clone="yk270b5" /clone_lib="yuj1 Kohara unpublished cDNA library" |

```

/dev_stage="varied"
/sex="hermaphrodite, male"
/tissue_type="whole animal"
BASE COUNT      145 a      65 c      55 g      88 t      7 others
ORIGIN

Query Match      6.0%; Score 37.4; DB 22; Length 360;
Best Local Similarity 63.6%; Pred. No. 1.6;
Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 535 cccaccttcctcctcagcaaccacactgtgagtgtgccaagtgcagcagaccttaaa 594
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 165 CTCACCCGACCCACCCACACCATCGAGTATTATTAANAATTTGATTTTAA 224

QY 595 aaaaaaagaaaaaacaacaaaaa 622
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 225 AAAAAAAAAAAAAAAAAAAAAA 252

RESULT 10
AA837213/c      289 bp      mRNA      EST      31-MAY-1998
LOCUS      od24f11.s1 NCI_CGAP_GCB1 Homo sapiens cDNA clone IMAGE:1368909,
DEFINITION      mRNA sequence.
ACCESSION      AA837213
NID      q2912412
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 (bases 1 to 299)
AUTHORS      NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE      National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
JOURNAL      Tumor Gene Index
COMMENT      Unpublished (1997)

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert.Strausberg@nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,
Ph.D., Gerald Marti, M.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/ILMIL at:
www-bio.11nl.gov/bdrp/image/image.html

Insert Length: 721 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 289.
Location/Qualifiers
1..299
/organism="Homo sapiens"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from human tonsillar cells enriched for
germinal center B cells by flow sorting (CD20+, IGD-),
provided by Dr. Louis M. Staudt (NCI), Dr. David Allman
(NCI) and Dr. Gerald Marti (CEBR). cDNA synthesis was
primed with a Not I - oligo(dT) primer
15'-TGTTCAACATCTAAGTGGAGCGGCCCTCTTTTCTTTTCTTTTCTTTT-
3'. Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="IMAGE:1368909"
/clone_11b="NCI_CGAP_GCB1"

```

```

/tissue_type="germinal center B cell"
/lab_host="DH10B"
BASE COUNT      66 a      78 c      65 g      90 t
ORIGIN

Query Match      5.9%; Score 37; DB 27; Length 299;
Best Local Similarity 67.5%; Pred. No. 2;
Matches 52; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

QY 546 cctacgaaccacactgtgagtgtgccaagtgcagcagaccttaaaaaaagaa 605
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 81 CCTACGACTCCCATCTGTCGCGTGAATAATAGTCTTCTCCTCCCAAAAAAAAA 22

QY 606 aaaaaaacaacaaaaa 622
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 21 AAAAAAAAAAAAAAAAAAAAA 5

RESULT 11
M91492/c      219 bp      mRNA      EST      29-OCT-1992
LOCUS      HUMRTPEAK Homo sapiens cDNA.
DEFINITION      M91492
ACCESSION      M91492
NID      q275908
KEYWORDS      EST.
SOURCE      Human library-Subtracted human retinal pigment epithelium (RPE).
ORGANISM      Homo sapiens
Eukaryota; Animalia; Chordata; Vertebrata; Mammalia; Theria;
Eutheria; Primates; Haplorhini; Catarrhini; Homnidae.
REFERENCE      1 (bases 1 to 219)
AUTHORS      Gieser, L. and Swaroop, A.
TITLE      Expressed sequence tags and chromosomal localization of cDNA clones
JOURNAL      from a subtracted retinal pigment epithelium library
MEDLINE      Genomics 13 (3), 873-876 (1992)
COMMENT      92347897

Contact: Swaroop, A. and Gieser, L.
Department of Ophthalmology
Kellogg Eye Center
University of Michigan, Ann Arbor, MI 48105.
Location/Qualifiers
1..219
/organism="Homo sapiens"
BASE COUNT      36 a      52 c      54 g      77 t
ORIGIN

Query Match      5.9%; Score 36.8; DB 10; Length 219;
Best Local Similarity 63.6%; Pred. No. 2.1;
Matches 56; Conservative 0; Mismatches 32; Indels 0; Gaps 0;

QY 535 cccaccttcctcctcagcaaccacactgtgagtgtgccaagtgcagcagaccttaaa 594
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 97 CCGCCACACTGTGAGGTCCACACGAGCTATTAAAGGAGATGTACTGCCAAAAAAA 38

QY 595 aaaaaaagaaaaaacaacaaaaa 622
      |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
Db 37 AAAAAAAAAAAAAAAAAAAAAA 10

RESULT 12
W37991/c      462 bp      mRNA      EST      15-MAY-1996
LOCUS      zc13609.s1 Soares parathyroid tumor NDHPA Homo sapiens cDNA clone
DEFINITION      322216 3', mRNA sequence.
ACCESSION      W37991
NID      g1319604
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
Eukaryota; mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

| REFERENCE | 1 (bases 1 to 462) |
|-----------|-----------------------|
| AUTHORS | Hillier, L., Clark, N |

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 14, 1999, 03:35:05 ; Search time 2254.51 Seconds
(without alignments)
1999.419 Million cell updates/sec

Title: US-09-212-270-10

Perfect score: 1260
Sequence: 1 ctctctagagggactgcgaac.....tgcgtgtgtgtatagatgaggg 1260

Scoring table: IDENTITY_NUC

Searched: 808301 seqs, 1788773984 residues

Database :

GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl1:*
9: gb_pl2:*
10: gb_pr1:*
11: gb_pr2:*
12: gb_pr3:*
13: gb_ro:*
14: gb_sc:*
15: gb_sy:*
16: gb_un:*
17: gb_vl:*
18: gb_htg:*
19: em_da:*
20: em_fun:*
21: em_hum1:*
22: em_hum2:*
23: em_in:*
24: em_om:*
25: em_or:*
26: em_ov:*
27: em_pat:*
28: em_ph:*
29: em_pl:*
30: em_ro:*
31: em_sy:*
32: em_un:*
33: em_vl:*
34: em_htg:*
35: em_sts:*
36: gb_ba1:*
37: gb_ba2:*
38: gb_pl1:*
39: gb_pl2:*
40: gb_pr1:*
41: gb_pr2:*
42: gb_pr3:*
43: gb_sts:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----|-------------|
| ----- | | | | | | |

| | | | | | | |
|----|--------|------|--------|----|------------|--------------------|
| 1 | 1233.2 | 97.9 | 1348 | 12 | AF046888 | AF046888 Homo sapi |
| 2 | 1233.2 | 97.9 | 1348 | 42 | AF046888 | AF046888 Homo sapi |
| 3 | 300 | 23.8 | 405 | 43 | G37120 | G37120 SHGC-56910 |
| 4 | 110 | 8.7 | 288 | 10 | HS40D10F | 260980 H.sapiens C |
| 5 | 110 | 8.7 | 288 | 40 | HS40D10F | 260980 H.sapiens C |
| 6 | 68.2 | 5.4 | 7218 | 6 | 166494 | 166494 Sequence 14 |
| 7 | 44.6 | 3.5 | 7218 | 6 | 166494 | 166494 Sequence 14 |
| 8 | 44.4 | 3.5 | 223 | 10 | HS40D10R | 260981 H.sapiens C |
| 9 | 44.4 | 3.5 | 223 | 10 | HS40D10R | 260981 H.sapiens C |
| 10 | 42.4 | 3.4 | 256404 | 18 | AC005071 | AC005071 *** SEQUE |
| 11 | 41.6 | 3.3 | 116917 | 12 | AC003024 | AC003024 Human Chr |
| 12 | 41.6 | 3.3 | 116917 | 42 | AC003024 | AC003024 Human Chr |
| 13 | 41.2 | 3.3 | 33233 | 10 | HSU51280 | U51280 Human chrom |
| 14 | 41.2 | 3.3 | 53522 | 11 | HUMPRK1GEN | L39891 Homo sapien |
| 15 | 41.2 | 3.3 | 200906 | 18 | HSU51280 | U51280 Human chrom |
| 16 | 41.2 | 3.3 | 33233 | 40 | HSU51280 | U51280 Human chrom |
| 17 | 41.2 | 3.3 | 53522 | 41 | HUMPRK1GEN | L39891 Homo sapien |
| 18 | 41.2 | 3.3 | 111863 | 18 | HS73E16 | 295330 Human DNA s |
| 19 | 40.6 | 3.2 | 37005 | 11 | AC005239 | AC005239 Homo sapi |
| 20 | 40.6 | 3.2 | 135046 | 11 | HSU91325 | U91325 Human chrom |
| 21 | 40.6 | 3.2 | 37005 | 41 | AC005239 | AC005239 Homo sapi |
| 22 | 40.6 | 3.2 | 135046 | 41 | HSU91325 | U91325 Human chrom |
| 23 | 40 | 3.2 | 317311 | 18 | AC004085 | AC004085 *** SEQUE |
| 24 | 39.8 | 3.2 | 4433 | 40 | HUMINCPSS | M19171 Human cysle |
| 25 | 39.8 | 3.2 | 4433 | 40 | HUMINCPSS | M19171 Human cysle |
| 26 | 39.6 | 3.1 | 57493 | 11 | AC004505 | AC004505 Homo sapi |
| 27 | 39.6 | 3.1 | 202004 | 11 | HUAF001549 | AF001549 Human Chr |
| 28 | 39.6 | 3.1 | 57493 | 41 | AC004505 | AC004505 Homo sapi |
| 29 | 39.6 | 3.1 | 202004 | 41 | HUAF001549 | AF001549 Human Chr |
| 30 | 39.4 | 3.1 | 47991 | 12 | AC005775 | AC005775 Homo sapi |
| 31 | 39.4 | 3.1 | 112444 | 18 | AC003022 | AC003022 *** SEQUE |
| 32 | 39.4 | 3.1 | 47991 | 42 | AC005775 | AC005775 Homo sapi |
| 33 | 39.2 | 3.1 | 102703 | 11 | AC004910 | AC004910 Homo sapi |
| 34 | 39.2 | 3.1 | 112846 | 11 | AC004991 | AC004991 Homo sapi |
| 35 | 39.2 | 3.1 | 151630 | 11 | AC005154 | AC005154 Homo sapi |
| 36 | 39.2 | 3.1 | 194020 | 11 | HUAC004125 | AC004125 Homo sapi |
| 37 | 39.2 | 3.1 | 102703 | 41 | AC004910 | AC004910 Homo sapi |
| 38 | 39.2 | 3.1 | 112846 | 41 | AC004991 | AC004991 Homo sapi |
| 39 | 39.2 | 3.1 | 151630 | 41 | AC005154 | AC005154 Homo sapi |
| 40 | 39.2 | 3.1 | 194020 | 41 | HUAC004125 | AC004125 Homo sapi |
| 41 | 39 | 3.1 | 20521 | 11 | HSMYCENH2 | 246773 H.sapiens D |
| 42 | 39 | 3.1 | 77236 | 18 | AC004581 | AC004581 *** SEQUE |
| 43 | 39 | 3.1 | 177236 | 18 | AC005143 | AC005143 *** SEQUE |
| 44 | 39 | 3.1 | 20521 | 41 | HSMYCENH2 | 246773 H.sapiens D |
| 45 | 38.8 | 3.1 | 101072 | 42 | AC005752 | AC005752 Homo sapi |

ALIGNMENTS

| | | | | | |
|------------|---|---------|------|-----------|-------------|
| RESULT 1 | AF046888 | 1348 bp | mRNA | PRI | 26-SEP-1998 |
| LOCUS | AF046888 | | | | |
| DEFINITION | Homo sapiens proliferation inducing ligand APRIL mRNA, complete cds. | | | | |
| ACCESSION | AF046888 | | | | |
| NID | G3650491 | | | | |
| KEYWORDS | human. | | | | |
| SOURCE | human. | | | | |
| ORGANISM | Homo sapiens | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo. | | | | |
| AUTHORS | Hahne,M., Kataoka,T., Schreuter,M., Hofmann,K., Immler,M., Bodmer,J.-L., Schneider,P., Bonnard,T., Holler,N., French,L.E., Sordal,B., Rimoldi,D. and Tschopp,J. | | | | |
| TITLE | APRIL, a new ligand of the tumor necrosis factor family, stimulates tumor cell growth | | | | |
| JOURNAL | J. Exp. Med. | 188 | (6), | 1185-1190 | (1998) |
| MEDLINE | 98416181 | | | | |
| REFERENCE | 2 (bases 1 to 1348) | | | | |
| AUTHORS | Hahne,M. and Tschopp,J. | | | | |
| TITLE | Direct Submission | | | | |


```

CDs      282..1034
         /note="TNF ligand family"
         /codon_start=1
         /product="proliferation inducing ligand APRIL"
         /db_xref="PID:93650493"
         /translation="MPASSPFLAEPKPPNGGPPVREPALSVLMLSWGALGAVAC
         AMALLOOTERLOSILREVSRLOGTGGSPONGGPMOSIPEOSDPALEWMEGERSRK
         RAVLTKOKKHOHSHLVLPYINATSKDSVDYEVAMQPIRLRGRLQAGYVRIODA
         GYLLISQVLFODVIFIMQVSRSGQOETLFRICIRMSHPHDPNANSCYSAQVFI
         LHQDILSVILIRARAKMLSPHGTFLGFKL"
         366..428
         misc_feature
         /note="transmembrane-region site"
BASE COUNT 293 a 377 c 374 g 304 t
ORIGIN
Query Match      97.9%; Score 1233.2; DB 42; Length 1348;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 1257; Conservative 0; Mismatches 3; Indels 2; Gaps 2;

QY      1 cttctagagagagctggaacctatctcctgagctgagagagtgagaggtcacaag 60
DB      10 cttctagagagagctggaacctatctcctgagctgagagagtgagaggtcacaag 69
QY      61 ggaagcctgagagagagagagagtgagagagagagagagagagagagagagagag 120
DB      70 gcaacgctgagagagagagagagagagagagagagagagagagagagagagagag 129
QY      121 tctcctcctcttcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 180
DB      130 tctcctcctcttcttcttcttcttcttcttcttcttcttcttcttcttcttctt 189
QY      181 cctctgcaacctgagcagcagcagcagcagcagcagcagcagcagcagcagcag 240
DB      190 cctctgcaacctgagcagcagcagcagcagcagcagcagcagcagcagcagcag 249
QY      241 gaaacagagcgtgtgagagagagagagagagagagagagagagagagagagagag 300
DB      250 gaaacagagcgtgtgagagagagagagagagagagagagagagagagagagagag 309
QY      301 ccccccaga-ggcccagagcaacatgagggggccagctcagagagagagagagag 359
DB      310 ccccccaga-ggcccagagcaacatgagggggccagctcagagagagagagagag 369
QY      360 cccctggttggatgtgaggggagcagctcagggcggtgtgtgtgtgtgtgtgt 418
DB      370 cccctggttggatgtgaggggagcagctcagggcggtgtgtgtgtgtgtgtgtgt 429
QY      419 cccaacaaacagagctgagagagagagagagagagagagagagagagagagagag 478
DB      430 cccaacaaacagagctgagagagagagagagagagagagagagagagagagagag 489
QY      479 gcccctcccaagatgagagagagatccctgagagagagagagagagagagagagag 538
DB      490 gcccctcccaagatgagagagagatccctgagagagagagagagagagagagagag 549
QY      539 cccctggaagccttggagagatgagagagagagagagagagagagagagagagagag 598
DB      550 cccctggaagccttggagagatgagagagagagagagagagagagagagagagagag 609
QY      599 aagaagaagaagcagcagctgtcctcagcagctgttcccattaagcagcagcagagatg 658
DB      610 aagaagaagaagcagcagctgtcctcagcagctgttcccattaagcagcagcagagatg 669
QY      659 actcgaatgtgagagagagatgtgagagagagagagagagagagagagagagagagag 718
DB      670 actcgaatgtgagagagagatgtgagagagagagagagagagagagagagagagagag 729
QY      719 cccaagagataggtgtcagagatccagagatgtgtgtgtgtgtgtgtgtgtgtgtgt 778
DB      730 cccaagagataggtgtcagagatccagagatgtgtgtgtgtgtgtgtgtgtgtgtgt 789
QY      779 tgttcaagaagcgtgacttcaacatgaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 838

```

```

DB      790 TGTTTAAGACGTGACTTTCCACCATGCTGTCAGTGTCTCTCAGAGACCAAGAGAGCC 849
QY      839 aggaagctatctccagatgtatagaagatagcctcccaaccgagacgggctacaaca 898
DB      850 AGGAGACTATTCGATGATATAGAAGATAGCCCTCCACCCGGAGACGGGCTTACACA 909
QY      899 gctgcataagcagagtgcttcccatcattcaaccaagagagatctgagtgctcaatc 958
DB      910 GCTGCTATGACGCGAGGTGCTTCCATTACACCAAGGAGATATTGAGTGTCTATATTC 969
QY      959 cccggagcaaggcggaacctatcctcctccatgagagagagagagagagagagagagag 1018
DB      970 CCGGGAGAGGGGAGAACTTACCTCTCTCCACATGGAACCTTCGTGGGGTGTGTAAC 1029
QY      1019 tctgattgttataaagtgctcccaagcttgaagacagagtggtgtacatactgga 1078
DB      1030 TGTGATTGTGTTATATAAAGTGGCTCCAGCTTGAAGACAGAGGTGGGTACTGTGA 1089
QY      1079 gaaagcaagagctgagatataaagagagagagagagagagagagagagagagagagag 1138
DB      1090 GACAGCCAAGAGCTGATATATAAGAGAGGGAATGTCCAGGACAGAGGCATCTTCCCT 1149
QY      1139 ggggtggtcccgcttcccaacttcccttcttcttcccaaccctagagcttgaatt 1198
DB      1150 GGGTGTGGCTCCCGCTTCCCTCACTTTCCCTTTCATTCCTCCACCCCTGAGCTTGAATT 1209
QY      1199 taagatatactgcttcttcccaagagagagagagagagagagagagagagagagagag 1258
DB      1210 TACGATATCTTCTCTCTTCTTCCCATGAGAGCTCCGAATTCTGCTGTGTATGATGAG 1269
QY      1259 gg 1260
DB      1270 GG 1271

```

```

RESULT      3
G37120/c    405 bp    DNA    STS    30-MAR-1998
LOCUS      SHGC-56910 Human Homo sapiens STS genomic, sequence tagged site.
DEFINITION G37120
ACCESSION  G2996771
NID        92996771
KEYWORDS   STS.
SOURCE      human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
            Primates; Catarrhini; Hominoidea; Homo.
REFERENCE   1 (bases 1 to 405)
            Myers,R.M.
            Human STS (1997)
            Unpublished (1997)
COMMENT
Contact: Richard M. Myers
Stanford Human Genome Center (SHGC)
Stanford University School of Medicine
Department of Genetics, M-344, Stanford, CA 94305, USA
Tel: 4157259687
Fax: 4157259689
Email: myers@shgc.stanford.edu
Primer A: ACACACACGCAAGATTCGG
Primer B: AGGAGAGGAGATGTGACGA
STS size: 149
PCR Profile:
Initial incubation: 95 degrees C for 10 minutes
Denaturation:      94 degrees C for 30 seconds
Annealing:          60 degrees C for 30 seconds
Polymerization:    72 degrees C for 23 seconds
PCR Cycles:         30
Thermal Cycler:     Perkin Elmer 9700
Template:            25 ng
Primer:              each 1 uM

```

dNTPs: each 200 uM
 AmpliTaq Gold Polymerase: 0.07 units/ul
 Total Vol: 5 ul

Buffer: MgCl₂: 2.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3

Prepared with primer pairs derived from AA057069 -- Unigene.

FEATURES
 source
 Location/Qualifiers
 1..405
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /map="17"
 /clone_11b="Human"

SRS
 primer_bind
 primer_bind complement(195..214)
 BASE COUNT 107 a 110 c 99 g 88 t 1 others
 ORIGIN

Query Match 23.8%; Score 300; DB 43; Length 405;
 Best Local Similarity 96.8%; Pred. No. 2e-77;
 Matches 338; Conservative 0; Mismatches 6; Indels 5; Gaps 3;

QY 917 tcttcattacacacagggatattctgagtg-cataattcccggaaggc--ga 973
 |||||||
 Db 405 TCTTCATTACACCAAGNGATTTCTGAGTGCATTAATCCCGGCAAGGCGCGA 346

QY 974 aactaacctctctc--acatgaacctctctggttggttggaactgtgattgtta 1031
 |||||||
 Db 345 AACTTAACCTCTCTCCACATGSAACCTTCTGCGTTGTGAACTGTGATTGTGTA 286

QY 1032 taaaagtggctccacagcttgaagaccagggtgtgtacactgagaaagcagaagc 1091
 |||||||
 Db 285 TAAAGATGGCTCCACACTTGAAGACAGGAGGTGGTACATACGAGACAGCAAGAGC 226

QY 1092 ttagtatataaaggaggggaatgtgcaggaacagagagcatctctggttggtctcc 1151
 |||||||
 Db 225 TGAATATATAAAGAGAGGAGATGTGACAGACAGAGCGCTCTCTGCGTTGGCTCC 166

QY 1152 cgttcctcacttttccctttcattccacacccctagactttgatttaagagatcttg 1211
 |||||||
 Db 165 CGTTCCTCATTCTTCCCTTTTCATTCACCCCTTAAGACTTTGATTTACGGAATCTTG 106

QY 1212 ctctgttcccatgagagctccgaattcttgcgtgtgtgtagatgaagg 1260
 |||||||
 Db 105 CTCTGTCTCCCATGAGAGCTCCGAATCTTGCGTGTGTAGATGAGAGG 57

RESULT 4
 LOCUS HS40D10F 288 bp DNA PRI 22-OCT-1995
 DEFINITION H.sapiens Cpg island DNA genomic MseI fragment, clone 40d10,
 forward read cpg40d10.ftla.
 ACCESSION 260980
 MID g1033358
 KEYWORDS Cpg island; genomic MseI fragment.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Eukaryotes; Metazoa; Chordata;
 Vertebrata; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 288)
 AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
 TITLE Direct Submission
 JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
 2 (bases 1 to 288)
 REFERENCE Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
 AUTHORS Purification of cpg islands using a methylated DNA binding column
 TITLE

JOURNAL Nature Genet. 6 (3), 236-244 (1994)
 MEDLINE 94282070
 COMMENT Vector: pGEM-52f(-)
 Clones are available from the UK MRC Human Genome Mapping Project
 Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
 http://www.hgmp.mrc.ac.uk/ for details
 or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES
 source
 Location/Qualifiers
 1..288
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /dev_stage="adult"
 /sex="male"
 /tissue_type="blood"
 /clone_11b="CGI-1"
 /clone="40d10"

BASE COUNT 77 a 68 c 85 g 57 t 1 others
 ORIGIN

Query Match 8.7%; Score 110; DB 10; Length 288;
 Best Local Similarity 100.0%; Pred. No. 1e-21;
 Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 gatgactcagatgtgacagagtgatgtgcaaccagctcttagcggtggaagagccta 714
 |||||||
 Db 142 GATGACTCCGATGTGACAGAGTGATGTGCAACAGCAGCTTATGGCGTGGAGAGGCTTA 201

QY 715 caggcccaagatatgtgtctcgatccagatcagatgctgagattatctgt 764
 |||||||
 Db 202 CAGGCCCAAGATATGATGTGCCAATCCAGAGTGTGAGATTATCTGCA 251

RESULT 5
 LOCUS HS40D10F 288 bp DNA PRI 22-OCT-1995
 DEFINITION H.sapiens Cpg island DNA genomic MseI fragment, clone 40d10,
 forward read cpg40d10.ftla.
 ACCESSION 260980
 MID g1033358
 KEYWORDS Cpg island; genomic MseI fragment.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Eukaryotes; Metazoa; Chordata;
 Vertebrata; Euthera; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 288)
 AUTHORS MacDonald,M., Huckle,E., Wilkinson,P. and Micklem,G.
 TITLE Direct Submission
 JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire,
 CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
 2 (bases 1 to 288)
 REFERENCE Cross,S.H., Charlton,J.A., Nan,X. and Bird,A.P.
 AUTHORS Purification of Cpg islands using a methylated DNA binding column
 JOURNAL Nature Genet. 6 (3), 236-244 (1994)
 MEDLINE 94282070
 COMMENT Vector: pGEM-52f(-)
 Clones are available from the UK MRC Human Genome Mapping Project
 Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL:
 http://www.hgmp.mrc.ac.uk/ for details
 or contact: biohelp@hgmp.mrc.ac.uk.

FEATURES
 source

1..288
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /dev_stage="adult"
 /sex="male"
 /tissue_type="blood"
 /clone_11b="CGI-1"
 /clone="40d10"
 BASE COUNT 77 a 68 c 85 g 57 t 1 others
 ORIGIN

Query Match 8.7%; Score 110; DB 40; Length 288;
Best Local Similarity 100.0%; Pred. No. 1e-21;
Matches 110; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 655 gatgactccgagtgtacagaggtatgtatgcaacacgctcttaagcgttggagagagccta 714
DB 142 GATGACTCCGATGTGACAGAGTGTATGGCAACACACTCTTACGCGTGGAGAGGCCTA 201
QY 715 caggcccaagatattgtgttcggaatccagatgtctgtgattcttgcct 764
DB 202 CAGGCCCAAGATATGTTGTCGATCCAGATCGAGATCTGAGTTATCTGCT 251

RESULT 6
LOCUS 166494 7218 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
NID 92724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner, F., Scheiflinger, F. and Falkner, F., Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
source 1. 7218
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 5.4%; Score 68.2; DB 6; Length 7218;
Best Local Similarity 11.8%; Pred. No. 2.5e-09;
Matches 34; Conservative 155; Mismatches 98; Indels 0; Gaps 0;

QY 19 acctaatctctcgtagagctgaggagggtgaggtctcaaggcaagctggcccaaga 78
DB 991 AACATATTTTCTCTGTTGTCATACGCTCACAGATTAATTCGAGCTTGCTGCAGGT 1050
QY 79 cggagtgccagagagcactaacagacacctgtgcttctctctctctctcttatt 138
DB 1051 CGAGGAGCTTGCGATYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1110
QY 139 ttcagttccttttattctctctcgtcgtacacacctctctctctctgcacacctgcc 198
DB 1111 YY 1170
QY 199 gtaccttaccgcccgcacacctctctgtctacccaccttgaacacacagacgttggc 258
DB 1171 YY 1230
QY 259 aggggtcccccagctatcgcacacctcattctctctgtcgtacgcccc 305
DB 1231 YY 1277

RESULT 7
LOCUS 166494/c 7218 bp DNA PAT 23-DEC-1997
DEFINITION Sequence 14 from patent US 5670367.
ACCESSION 166494
NID 92724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE 1 (bases 1 to 7218)
AUTHORS Dörner, F., Scheiflinger, F. and Falkner, F., Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;

FEATURES
source Location/Qualifiers
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN

Query Match 3.5%; Score 44.6; DB 6; Length 7218;
Best Local Similarity 2.1%; Pred. No. 0.021;
Matches 8; Conservative 216; Mismatches 155; Indels 0; Gaps 0;

QY 363 tctgtgtgattggggggcagctcgtggggcggttggtgcatggttcagtcgaacca 422
DB 1442 TTTCGTACRR 1383
QY 423 acaacagagctgcagagcctcagagagaggtgagcgcgtgcagggagacagagccc 482
DB 1382 RRR 1323
QY 483 ctcccagaatggggaaggtatccctgtgcagagctcctccgagacagagttcgcacct 542
DB 1322 RRR 1263
QY 543 ggaagcctggagagatggggagagatcccgaaagagagcagtgctcaccacaaaca 602
DB 1262 RRR 1203
QY 603 gaagaagcagacactgtctcgtcacctggttccattacgccaccctcaagatgactc 662
DB 1202 RRR 1143
QY 663 cgaatgtacagaggtgtatgtgcacacagctctagcgttggagagagcctacagccca 722
DB 1142 RRR 1083
QY 723 aggataagtgctcgcaatc 741
DB 1082 RRR 1064

RESULT 8
LOCUS HS40D10R/c 223 bp DNA PRI 22-OCT-1995
DEFINITION H.sapiens Cpg island DNA genomic MseI fragment, clone 40d10, reverse read cpg40d10.rta.
ACCESSION 260981
NID 91033359
KEYWORDS Cpg island; genomic MseI fragment.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotae; Mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Eutheria; Primates; Carnivora; Homnidae; Homo.

REFERENCE 1 (bases 1 to 223)
AUTHORS MacDonald, M., Huckle, E., Wilkinson, P. and Mickle, G.
TITLE Direct Submision
JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire, CB10 1RQ, England. E-mail contact: humgen@esanger.ac.uk

REFERENCE 2 (bases 1 to 223)
AUTHORS Cross, S.H., Charlton, J.A., Nan, X. and Bird, A.P.
TITLE Purification of cpg islands using a methylated DNA binding column
JOURNAL Nature Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070
COMMENT Vector: pGEM-5zf(-)
Clones are available from the UK MRC Human Genome Mapping Project Resource Centre, Hinxton, Cambridgeshire CB10 1RQ, UK. See URL: <http://www.hgmp.mrc.ac.uk/> for details or contact: biohelp@hgmp.mrc.ac.uk.
Location/Qualifiers
1. 223
/organism="Homo sapiens"
/db_xref="taxon:9606"
/dev_stage="adult"
/sex="male"

| BASE COUNT | ORIGIN | 51 a | 68 c | 44 g | 54 t | 6 others |
|------------|--------|----------------------|------|------|------|----------|
| | | /tissue_type="blood" | | | | |
| | | /clone_1b="CGI-1" | | | | |
| | | /clone="40d10" | | | | |

Query Match * 3.58; Score 44.4; DB 10; Length 223;
Best Local Similarity 94.48; Pred. No. 0.018;
Matches 67; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Oy 909 cgcaggtgcttcattacaccaagggatatt-ctgaagtgtcataaa-ttccccggcgca 966
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 71 CTCAGGTGCTTCATTACACCAGAAGGATATTNCTGAAGTGCATAATTTCNCGGGCCA 12

| | | | |
|----|-----|-------------|-----|
| QY | 967 | agggcggaact | 977 |
| | | | |
| Db | 11 | AGGCGGAAC | 1 |

| | | | | |
|------------|--|--------|-----|-------------|
| RESULT | 9 | | | |
| HS40D10R/c | | | | |
| LOCUS | HS40D10R | 223 bp | DNA | PRI |
| DEFINITION | H.sapiens CpG island DNA genomic MseI fragment, clone 40d10, | | | |
| | reverse read cpg40d10.tla. | | | |
| ACCESSION | 760091 | | | 22-OCT-1995 |

| | |
|----------|-----------------------------------|
| IND | 91033333 |
| KEYWORDS | Cpg island; genomic MseI fragment |
| SOURCE | human. |
| ORGANISM | Homo sapiens |

REFERENCE
AUTHORS
TITLE
Direct Submission

Vertebrata; Eutheria; Primates; Catarrhini; Hominiidae; Homo
1 (bases 1 to 223)
Macdonald, M., Huckle, E., Wilkinson, P. and Micklem, G.

JOURNAL Submitted (16-OCT-1995) The Sanger Centre, Hinxton, Cambridgeshire
 CB10 1RQ, England. E-mail contact: humquerry@sanger.ac.uk
 REFERENCE 2 (bases 1 to 223)

TITLE Purification of Cpg islands using a methylated DNA binding column
JOURNAL Nature Genet. 6 (3), 236-244 (1994)
MEDLINE 94282070

COMMENT
vector: pGEM-521 (-)
Clones are available from the UK MRC Human Genome Mapping Project
Resource Centre, Hinxton, Cambridgeshire CB10 1NQ, UK. See URL:
<http://www.hgmp.mrc.ac.uk/> for details
or contact: blonhelp@hgmp.mrc.ac.uk.

| FEATURES | Location/Qualifiers |
|----------|---------------------|
| source | 1. .223 |

| BASE COUNT | ORIGIN |
|------------|----------|
| 51 a | 68 c |
| 44 g | 54 t |
| | 6 others |

| | | | | |
|-----------------------|-----------------|------------------|-----------|-------------|
| Query Match | 3.5% | Score 44.4; | DB 40; | Length 223; |
| Best Local Similarity | 94.4%; | Pred. No. 0.018; | | |
| Matches 67; | Conservative 0; | Mismatches 2; | Indels 2; | Gaps 2 |

QY 909 cgcaggtgtcttcattacaccagggaatat-ctgagtgtcaata-ttccccggca 96
| | | | |
Db 71 CTCAGGTGTCCTCCATTACACCAGGGCATATNCTGAGTGTCAATAATTCNCCGGCA 12

| | | | |
|----|-----|-------------|-----|
| Qy | 967 | agggcgaaact | 977 |
| | | | |
| Db | 11 | AGGGCGAAACT | 1 |

| | | | |
|------------|--|--------------------------------|-----------------------------------|
| RESULT | 10 | | |
| LOCUS | AC005071 | | |
| DEFINITION | AC005071 256404 bp DNA | HTG | 12-JUN-1998 |
| ACCESSION | AC005071 | *** SEQUENCING IN PROGRESS *** | Homo sapiens clone RG161A02; HTGS |
| NID | g3212920 | phase 1, 35 unordered pieces. | |
| KEYWORDS | HTG; HTGS_PHASE1. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryotes: Metazoa; Chordata; Vertebrata; Mammalia; Euthera; | | |
| AUTHORS | Primates; Catarrhini; Homnidae; Homo. | | |
| TITLE | 1 (bases 1 to 256404) | | |
| JOURNAL | Waterston, R.H. | | |
| REFERENCE | The sequence of Homo sapiens clone | | |
| AUTHORS | Unpublished | | |
| TITLE | 2 (bases 1 to 256404) | | |
| JOURNAL | Waterston, R.H. | | |
| COMMENT | Direct Submission | | |
| | Submitted (12-JUN-1998) Genome Sequencing Center, Washington | | |
| | University School of Medicine, 4444 Forest Park Parkway, St. Louis | | |
| | MO 63108, USA | | |
| | *** | | |
| | *** | | |

* This sequence is unfinished. It consists of 35 contigs for
* this sequence is not known; their order in this record is
* arbitrary. In some cases, the exact lengths of the gaps
* between the contigs are also unknown; these gaps are presented
* as runs of N (a convenience only). When sequencing is complete
* the sequence data presented in this record will be replaced
* by a single finished sequence with the same accession number.

| | | | |
|---|-------|--------|-----------------------------|
| 1 | 1601 | 1600: | contig of 1600 bp in length |
| * | 1619 | 1618: | gap of unknown length |
| * | 3214 | 3213: | contig of 1595 bp in length |
| * | 3232 | 4831: | gap of unknown length |
| * | 4832 | 4831: | contig of 1600 bp in length |
| * | 4850 | 4849: | gap of unknown length |
| * | 6666 | 6665: | contig of 1816 bp in length |
| * | 6684 | 6683: | gap of unknown length |
| * | 8473 | 8472: | contig of 1789 bp in length |
| * | 8491 | 8490: | gap of unknown length |
| * | 10002 | 10001: | contig of 1511 bp in length |
| * | 10020 | 10019: | gap of unknown length |
| * | 11651 | 11650: | contig of 1631 bp in length |
| * | 11669 | 11668: | gap of unknown length |
| * | 13498 | 13497: | contig of 1829 bp in length |
| * | 13498 | 13515: | gap of unknown length |
| * | 13516 | 13515: | contig of 2429 bp in length |
| * | 15945 | 15962: | gap of unknown length |
| * | 15963 | 17496: | contig of 1534 bp in length |
| * | 17497 | 17514: | gap of unknown length |
| * | 17515 | 19092: | contig of 1578 bp in length |
| * | 18093 | 19110: | gap of unknown length |
| * | 19111 | 21089: | contig of 1979 bp in length |
| * | 21090 | 21107: | gap of unknown length |
| * | 21108 | 22787: | contig of 1660 bp in length |
| * | 22788 | 22805: | gap of unknown length |
| * | 22806 | 24419: | contig of 1614 bp in length |
| * | 24420 | 24437: | gap of unknown length |
| * | 24438 | 25332: | contig of 1455 bp in length |
| * | 25931 | 25930: | gap of unknown length |
| * | 25951 | 27541: | contig of 1591 bp in length |
| * | 27542 | 27559: | gap of unknown length |
| * | 27560 | 29631: | contig of 2272 bp in length |
| * | 28832 | 29649: | gap of unknown length |
| * | 29850 | 31952: | contig of 2103 bp in length |
| * | 31953 | 31970: | gap of unknown length |
| * | 31971 | 33770: | contig of 1800 bp in length |
| * | 33771 | 33788: | gap of unknown length |
| * | 33789 | 35495: | contig of 1707 bp in length |
| * | 35496 | 35513: | gap of unknown length |
| * | 35514 | 37491: | contig of 1978 bp in length |

[illegible]

| | | | |
|-----------------------|--|-------------------|--|
| REFERENCE | AUTHORS | TITLE | JOURNAL |
| 3 (bases 1 to 116917) | Evans,G.A., Athanasou,M., Basil,M., Bradbury,P., Bignac,S., Bunesler,R., Davis,C., English,C., Franklin,T.L., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Hinson,S., Narayanaswamy,U., Newton,J., O'Brien,K., Patel,P., Schageman,J., Schilling,P., Schultz,R., Syed,M., Valenzuela,D., Ward,T. and Wilson,R. | Direct Submission | Submitted (21-OCT-1997) Genome Science and Technology Center, University of Texas Southwestern Medical Center at Dallas, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA |
| 4 (bases 1 to 116917) | Evans,G.A., Athanasou,M., Aguayo,P., Armstrong,D., Basil,M., Buetcher,J., Bunesler,R., Card,P., desallibot,F., Dunn,J., English,C., Ethridge,S., Garner,H.R., Gee,V., Gordon,M., Gotway,G., Grant,O., Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T., McFarland,J., Newton,J., Osborn-Lawrence,S., Schageman,J., Schultz,R.A., Stinson,S., Syed,M. and Ward,T. | Direct Submission | Submitted (17-DEC-1997) Genome Science & Technology Center, University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd, Dallas, TX 75235-8591, USA |
| COMMENT | On Sep 16, 1998 this sequence version replaced 41:2695560. Further information regarding the map of this region or annotation of pDU41616 can be found at http://genetec.swmed.edu/chromos5.htm . IMPORTANT: This submission contains the entire insert of clone pDU41616. pDU41616 comes from the RPCI-3 PAC library constructed at the Roswell Park Cancer Institute by the Pletzer de Jong group. This clone has been finished according to strict quality criteria and attempts have been made to resolve all base calling problems such as compressions and repetitive elements. The expected phred/phrap calculated errors/10kb is 0.56. In addition, attempts have been made to assure over 99% of consensus base calls consist of either double-stranded coverage or 2 types of labeling chemistry on one strand. CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 15q26.1 Bloom's syndrome/DNA polymerase gamma region. This region is mapped between STS D15S171 and D15S652 MARKER CONFIRMATION: STS sequence confirmed; WI-6813, D15S653, SHGC-9392 MAPPED CLONE OVERLAP: pDU422e7. | | |
| FEATURES | Location/Qualifiers | | |
| Source | 1..116917 | | |
| repeat_region | /organism="Homo sapiens" | | |
| repeat_region | /db_xref="taxon:9606" | | |
| repeat_region | 5376..5521 | | |
| repeat_region | /rpt_family="MIR" | | |
| repeat_region | complement(8545..8834) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 9913..10219 | | |
| repeat_region | /rpt_family="MER2" | | |
| repeat_region | 11974..12151 | | |
| repeat_region | /rpt_family="MLT1" | | |
| repeat_region | 15289..16374 | | |
| repeat_region | /rpt_family="L1" | | |
| repeat_region | 17666..17816 | | |
| repeat_region | /rpt_family="L1" | | |
| repeat_region | 18471..18642 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(20327..20615) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(20947..21077) | | |
| repeat_region | /rpt_family="MIR" | | |
| repeat_region | 21210..21508 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 23183..23483 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(26548..27241) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(27278..27386) | | |
| repeat_region | /rpt_family="MLT1" | | |
| repeat_region | 30624..31046 | | |
| repeat_region | /rpt_family="MER2" | | |
| repeat_region | complement(32058..33119) | | |
| repeat_region | /rpt_family="MER43" | | |
| repeat_region | complement(32746..32980) | | |
| repeat_region | /rpt_family="MER43" | | |
| repeat_region | 33385..37165 | | |
| repeat_region | /rpt_family="L1" | | |
| repeat_region | 37367..38347 | | |
| repeat_region | /rpt_family="MER42" | | |
| repeat_region | complement(43991..44279) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(44423..44623) | | |
| repeat_region | /rpt_family="MIR" | | |
| repeat_region | complement(47160..47446) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(49663..49937) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 50664..50947 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(52207..52266) | | |
| repeat_region | /rpt_family="MER6" | | |
| repeat_region | complement(52470..52541) | | |
| repeat_region | /rpt_family="MER46" | | |
| repeat_region | complement(58313..58613) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 59242..59382 | | |
| repeat_region | /rpt_family="MER46" | | |
| repeat_region | complement(59243..59474) | | |
| repeat_region | /rpt_family="MER46" | | |
| repeat_region | 61236..61507 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 63584..63715 | | |
| repeat_region | /rpt_family="MER5" | | |
| repeat_region | complement(63584..63714) | | |
| repeat_region | /rpt_family="MER5" | | |
| repeat_region | complement(70158..70391) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 72078..72148 | | |
| repeat_region | /rpt_family="L1" | | |
| repeat_region | 72205..72466 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 74340..74493 | | |
| repeat_region | /rpt_family="MIR" | | |
| repeat_region | complement(76377..76469) | | |
| repeat_region | /rpt_family="MIR" | | |
| repeat_region | 77426..77717 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(82461..82745) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 84074..84249 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(85332..85612) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 88044..88333 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 90721..90757 | | |
| repeat_region | /rpt_family="MIR" | | |
| repeat_region | 93074..93360 | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | complement(94966..94986) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 98459..98706 | | |
| repeat_region | /rpt_family="L1" | | |
| repeat_region | complement(98873..99025) | | |
| repeat_region | /rpt_family="MIR" | | |
| repeat_region | complement(99895..100145) | | |
| repeat_region | /rpt_family="Alu" | | |
| repeat_region | 100174..100500 | | |

| | | | |
|--------|---|------|---|
| intron | /gene="PKD1" /number=3 19975. .20095 /gene="PKD1" /number=3 20096. .20167 /gene="PKD1" /number=4 20168. .20435 /gene="PKD1" /number=4 20436. .20605 /gene="PKD1" /number=5 20606. .21109 /gene="PKD1" /number=5 21110. .21490 /gene="PKD1" /note="putative" /number=6 21174. .21459 /gene="PKD1" /note="Cpg island; putative" 21491. .21608 /gene="PKD1" /number=6 21609. .21792 /gene="PKD1" /note="putative" /number=7 21793. .22227 /gene="PKD1" /number=7 22228. .22448 /gene="PKD1" /note="putative" /number=8 22449. .22636 /gene="PKD1" /number=8 22637. .22752 /gene="PKD1" /note="putative" /number=9 22753. .23117 /gene="PKD1" /number=9 23118. .23289 /gene="PKD1" /note="putative" /number=10 23290. .23655 /gene="PKD1" /number=10 23656. .23815 /gene="PKD1" /note="putative" /number=11 23816. .24966 /gene="PKD1" /number=11 24967. .25111 /gene="PKD1" /note="putative" /number=12 25112. .25988 /gene="PKD1" /number=12 25989. .26120 /gene="PKD1" /note="putative" /number=13 | exon | 26121. .26317 /gene="PKD1" /number=13 26318. .26493 /gene="PKD1" /note="putative" /number=14 26494. .26807 /gene="PKD1" /number=14 26808. .26941 /gene="PKD1" /note="putative" /number=15 26942. .27406 /gene="PKD1" /number=15 27407. .28254 /note="Cpg island; putative" 27407. .27525 /gene="PKD1" /note="putative" /number=16 27526. .27627 /gene="PKD1" /number=16 27628. .28683 /gene="PKD1" /note="putative" /number=17 28608. .28993 /note="Cpg island; putative" 28684. .29125 /gene="PKD1" /number=17 29126. .29875 /gene="PKD1" /note="putative" /number=18 29857. .30825 /note="Cpg island; putative" 29876. .30418 /gene="PKD1" /number=18 30419. .31024 /gene="PKD1" /note="putative" /number=19 31025. .31243 /gene="PKD1" /number=19 31244. .31393 /gene="PKD1" /note="putative" /number=20 31394. .32327 /gene="PKD1" /number=20 32328. .32471 /gene="PKD1" /note="putative" /number=21 32472. .32598 /gene="PKD1" /number=21 32599. .32878 /gene="PKD1" /note="putative" /number=22 32616. .33079 /note="Cpg island; putative" 32879. .32971 /gene="PKD1" |
|--------|---|------|---|

| | | | |
|---------------|--|---------------|--|
| exon | /number=22 32972..33185 /gene="PKD1" /note="putative" | exon | /note="Cpg island; putative" 41523..41733 /gene="PKD1" |
| intron | /number=23 33186..33250 /gene="PKD1" | intron | /number=33 41734..41823 /gene="PKD1" |
| exon | /number=23 33251..33410 /gene="PKD1" /note="putative" | exon | /number=33 41824..41950 /gene="PKD1" |
| intron | /number=24 33411..33847 /gene="PKD1" | intron | /number=34 41951..43609 /gene="PKD1" |
| exon | /number=24 33848..33952 /gene="PKD1" /note="putative" | repeat_region | /number=34 42463..42685 /rpt_family="Alu" |
| intron | /number=25 33953..36953 /gene="PKD1" | repeat_region | 43084..43321 /rpt_family="Alu" |
| exon | /number=25 36954..37098 /gene="PKD1" /note="putative" | exon | 43610..43726 /gene="PKD1" |
| misc_feature | /number=26 37007..37251 /note="Cpg island; putative" | intron | /number=35 43727..43816 /gene="PKD1" |
| intron | /number=26 37099..37700 /gene="PKD1" | exon | /number=35 43817..43866 /gene="PKD1" |
| exon | /number=26 37701..38330 /gene="PKD1" | intron | /number=36 43867..44090 /gene="PKD1" |
| intron | /number=27 38331..38623 /gene="PKD1" | exon | /number=36 44091..44275 /gene="PKD1" |
| exon | /number=27 38624..38780 /gene="PKD1" | intron | /number=37 44276..44352 /gene="PKD1" |
| intron | /number=28 38781..38960 /gene="PKD1" | exon | /number=37 44353..44446 /gene="PKD1" |
| exon | /number=28 38961..39213 /gene="PKD1" | intron | /number=38 44447..47383 /gene="PKD1" |
| intron | /number=29 39214..39337 /gene="PKD1" | repeat_region | 44778..45002 /rpt_family="Alu" |
| exon | /number=29 39338..39533 /gene="PKD1" | repeat_region | complement(46331..46602) /rpt_family="Alu" |
| intron | /number=30 39534..41027 /gene="PKD1" | exon | 47384..47502 /gene="PKD1" |
| repeat_region | /number=30 complement(39970..40228) /rpt_family="Alu" | intron | /number=39 47503..47580 /gene="PKD1" |
| repeat_region | 40311..40598 /rpt_family="Alu" | exon | /number=39 47581..47783 /gene="PKD1" |
| exon | 41028..41198 /gene="PKD1" | intron | /number=40 47784..47855 /gene="PKD1" |
| intron | /number=31 41199..41284 /gene="PKD1" | exon | /number=40 47856..48050 /gene="PKD1" |
| exon | /number=31 41285..41428 /gene="PKD1" | intron | /number=41 48051..48500 /gene="PKD1" |
| intron | /number=32 41429..41522 /gene="PKD1" | exon | /number=41 48501..48640 /gene="PKD1" |
| misc_feature | /number=32 41501..41730 | intron | /number=42 48641..49001 /gene="PKD1" |
| | | exon | /number=42 49002..49114 /gene="PKD1" |
| | | intron | /number=43 49115..49405 |

| Feature | Score | DB | Length |
|---------------------------|--|----|--------|
| exon | 49406.49547 | | |
| intron | 49548.49687 | | |
| exon | 49688.49813 | | |
| intron | 49814.49996 | | |
| misc_feature | 49898.50854 | | |
| exon | 49997.50171 | | |
| intron | 50172.50418 | | |
| exon | 50419.50709 | | |
| conflict | 50652.50653 | | |
| intron | 50710.50784 | | |
| exon | 50785.50919 | | |
| conflict | 50796.50797 | | |
| intron | 50920.51002 | | |
| exon | 51003.51308 | | |
| intron | 51309.51398 | | |
| exon | 51399.52883 | | |
| conflict | 51827.51828 | | |
| conflict | 52120 | | |
| polyA_signal | 52862.52867 | | |
| BASE COUNT | 8486 a 17664 c 15766 g 11606 t | | |
| ORIGIN | | | |
| Query Match | 3.38; Score 41.2; DB 11; Length 53522; | | |
| Best Local Similarity | 54.1; Pred. No. 0.25; | | |
| Matches 106; Conservative | 0; Mismatches 88; Indels 2; Gaps | | |

[illegible]

Db 136133 CTCCTTCCTTCCT 136120

Search completed: May 14, 1999, 03:36:08
Job time: 5787 sec

This Page Blank (uspto)

This Page Blank (uspto)

QY 130 c-ttttatttcgaagcttccttatttattctctctgtaaacactctctctctg 188
Db 137 ctttttttttttcaagttccttttttttttttttttttttttttttttttt
QY 189 aacctgcccgtaccttaaccgcccgcacactctctgtactaccactctgaaacac 248
Db 197 ACCACTGCCCGTACCCTTACCCTGCCGCCACACTCTGTCTACACACTCTGAAACAC 256
QY 249 agcttggcagggtt-cccacagctacgacagcctcatctctcttctgttagccccc 307
Db 257 AGCTTTGGCAGGGGTCCTCCCACTATGCAACCTCTCTTCTTGTGTAAGCCCCCA 316
QY 308 a-ggcctccaggaacaatgagggggccagctcagagagcgcacactcaagttgcccctg 366
Db 317 AGGGCTCCAGGCAACATGGGGGGCCAGTCAAGAGCGGACCTCAGTTGCCCTCTG 376
QY 367 gttgagttggggggcagctctggggggccgtgtgttg-catggttcagctgaacca 425
Db 377 GTTGAATTGGGGGGCAGCTCTGGGGGGCGGTGCTGTGCTCAAGGCTGTGACCAACA 436
QY 426 aacagagctgagagcctcagagagagagtgagccggtctcagggagagagggccctc 485
Db 437 AACAGAGCTGCAGAGCTCAGAGAGAGGTGAGCCGGCTGCAGAGAGAGAGGGCCCTC 496
QY 486 ccagaatggggaagggtatccctcgacagagctcccgagagcagagttccgattgccc 545
Db 497 CCAGAAATGGGGAAGGTTATCCCTGGCAGAGTCTCCCGAGCGAGATTCGATGCCCTGA 556
QY 546 agcctgggagagatgggagagatcccggaagaaagagagcagtgctacccaaacagaa 605
Db 557 AGCCCGGAGAGATGGGGAGAGATCCCGMAAGAGAGAGAGTCTAGCCAAAGAACAGNA 616
QY 606 gaagagagactgtgctcgacactgttcccaataagccacactcaagagagagactcga 665
Db 617 GAGAGAGACTGTGCTGCTGACACTGTTCCCTTAAAGCCACCTCCAAAGAGAGACTCCGA 676
QY 666 tgtgacagagtgatgtggaacacagctcttagcgttggagagagccctacagggccaa 725
Db 677 TGTGACAGAGGTGATGTGGCAACAGCTTATGGGCTGGGAGAGGCTAACAGGCCCAAG 736
QY 726 atatgtgtccgaatccagagatgctggagttatctgtctgtatagccaggtctgttca 785
Db 737 ATATGTGTCCGAATCCAGAGATGCTGGAGATTATCTGCTATACCGAGTCTTTTCA 796
QY 786 agaagctgttcacacagtggtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 845
Db 797 AGAGCTGACTTTACCAATGGGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 856
QY 846 tctattccgagtataaagaatgctccctccacccggagccgggacctacaacagctgta 905
Db 857 TCTATTCCGATGTATTAAGAGATATGCCCTCCACCCGGACCGGGCTTCAACAGCTGTGA 916
QY 906 tagcgcaagtgctctccattacacaaaggagatatctgtggtcacaattcccccggc 965
Db 917 TAGCCAGAGTGTCTTCAATTAACCAAGGGGATATCTGATGTCAATAATCCCCGGGC 976
QY 966 aaggcggaacttaacactctccacatggaacactcctggggttttggaaacttgatt 1025
Db 977 AAGGCGGAACCTTAACCTCTCTCCACATGAGACCTTCTCGGGGTTTGTGAACCTGTGAT 1036
QY 1026 gtttataaaagtgtctccagctctggaagacagaggttgggtataactctggaagacgc 1085
Db 1037 GTGTATATAAAAGTGGCTCCCGAGTTGGAAGACAGGGGTGTACTACTGAGACAGCC 1096
QY 1086 aagagctgagtataaagagagaggaatgtgcaagaaacagagcatctctgtggtttg 1145
Db 1097 AAGAGCTAGATATTAAGAGAGAGGAATGTGACAGAACAGAGGCTCTCTCGGGTTTG 1156
QY 1146 gctcccgcttcctcacttctccttcttcaatccacccctcagagattgatttgcgat 1205
Db 1157 GCTCCCTCTCTCACTTTTCCCTTTTCAATTCACACCCCTTGAATTTAGCGGAT 1216
QY 1206 atctgtctctgtctcccatgagctccgaattctgtggtgtatgagaggg 1260

Db 1217 AATCTGCTTCTGTTCCCATGAGCTCGAATCTTCTGCTGTGTGTGATGAGG 1271

|||||

RESULT 2

ID V00494

AC V00494 standard; cDNA; 1275 BP.

DE 14-APR-1998 (first entry)

DE Human tumour necrosis factor epsilon encoding cDNA.

KW Human; tumour necrosis factor delta; TNF-delta; TNF-epsilon; cancer;

KW tumour necrosis factor epsilon; T cell proliferation;

OS immune regulation; inflammatory response; ss.

OS Homo sapiens.

EH key

FT CDS

FT 2..508

FT Location/Qualifiers

PD MO9733902-A1.

PD 18-SEP-1997.

PF 14-MAR-1996; U03774.

PR 14-MAR-1996; WO-U03774.

PA (HOMA-) HUMAN GENOME SCI INC.

PI Gentz RL, Ni J, Yu G;

DR WPI: 97-470810/43.

DR P-PSDB: M37003.

PT Nucleic acid encoding human tumour necrosis factors delta and epsilon - useful for destroying cancer cells, or mediating T cell proliferation, immune regulation and inflammatory responses etc.

PT Claim 12; Fig 2; 83pp; English.

CC The present sequence encodes human tumour necrosis factor (TNF) epsilon. The TNF can be used to treat patients deficient in this factor (optionally by in vivo expression). It is a ligand for TNF and can destroy (by inducing apoptosis) some transformed cell lines (for tumour treatment); mediate cell activity and proliferation (including T cells to stimulate an immune response to viral, bacterial or parasitic infections, also to eliminate autoreactive T cells in e.g. type I diabetes), and is functionally linked as a primary mediator of immune regulation and the inflammatory response. A typical application of the protein, its agonists or antagonists is prevention of septic shock, inflammation, cerebral malaria, human immunodeficiency virus activation, graft-host reaction, bone resorption (osteoporosis), rheumatoid arthritis and cachexia, also to promote wound healing (by upregulating cell adhesion) and to regulate haematopoiesis. The protein can also be used to isolate its cognate receptors (or receptor-encoding genes). CC The cDNA can be used to produce recombinant proteins, and fragments of it to isolate full-length or related sequences, for diagnostic detection of mutations (indicative of disease or susceptibility) and for chromosome identification. Analysis of patient samples for presence of the new protein (e.g. using specific antibodies) can be used diagnostically.

CC Sequence 1275 BP: 283 A: 343 C: 356 G: 293 T:

SO

Query Match 53.3%; Score 671.2; DB 1; Length 1275;

Best Local Similarity 93.5%; Pred. No. 5; 9e-194;

Matches 742; Conservative 0; Mismatches 3; Indels 49; Gaps 2;

QY 468 gggagcagagagccctccagaaatgggaaaggtatccctgagagagttcccgagca 527

Db 1 GGGGACAGAGAGCCCTCCAGAAATGGGAGGGTATCTCCGCAAGATCCCGGAGCA 60

QY 528 gagttccgagtcctctggaagcctgggagaaatgggagagatcccggaagagagcagt 587

Db 61 GAGTTCGATCCCTCGAAGCTTGGAAGTGGGAGAGATCCCGAAGAAAGAGACAGT 120

QY 588 gctacccaaagagagagcagcactgtctcgtcagactggttcccatcagcagcc 647

Db 121 GCTACCCAAAGAGAGAA----- 141

QY 648 ctcaagagatctcagatgtgacagaggtgagtgtgcaacacagctttagcggtggag 707

Db 141 -----ATGACTTCGATGTGACAGAGGTGATGTGGCAACACACTCTTAGCGTGGAG 192

| | | | |
|----|------|---|------|
| QY | 708 | aggctctacagggcccaagatatggtgtccgaatccagagatgctgaggtattctctgtta | 767 |
| Db | 193 | AGGCTCTACAGGCCCCAAAGATATGGTGTCCCAATCCAGGATGCTGGAGTTATCTCTGTA | 252 |
| QY | 768 | tagccaagtccttggttcaaacgtgacttaccatgggtcaaggtgtgtctccggaag | 827 |
| Db | 253 | TAGCCAGTCTCTGTTTCAAGCTGACTTTCACCAATGGGTCAGGTGTCTCGGAAGG | 312 |
| QY | 828 | ccaaggaagcgagagaagactatcccatgtatataaagaatgtccctccaccgcgagcg | 887 |
| Db | 313 | CCAAGGAAGCGAGAGACTCTATTCGATGTATTAAGAAGATGCCCTCCACCGGACCG | 372 |
| QY | 888 | ggctctacaacagctgtctatagcgcaggtgtctctccattacaaccaaggagatcttag | 947 |
| Db | 373 | GGCTCTACAACAGCTGCTATAGCCAGGTGCTTCCATTACACCAAGGGGATATCTGAG | 432 |
| QY | 948 | tgtatataatcccgaggccaaggcggaacttaacctctctccaatggaactctctagg | 1007 |
| Db | 433 | TGTATATATTCCTCCGGGCAAGGGCGAAACTTAACCTCTCTCCACAGGAACCTTCTGGG | 492 |
| QY | 1008 | gttcgtgaactgtgactgtgttataaagtgctgccacagcttggaaagacaggttgg | 1067 |
| Db | 493 | GTTTGTGAACACTGTGATTGTGTATATAAAGTGGCTCCACGCTTGGAAACACAGGCTGG | 552 |
| QY | 1068 | tacatactggaagacagccaagagctgagatataaagaagaggaatgtgcagacaga | 1127 |
| Db | 553 | TACATACTGAGAGACAGCCAAAGAGCTGATATATAAAGAGAGGGAATGCGAGAAACA | 612 |
| QY | 1128 | ggactcttctcgtggttggctcccgctctcactttcccttccattccaccacctta | 1187 |
| Db | 613 | GGGCTGCTTCCTGGGTTGGCTCCCGCTCCCTCACTTTCCCTTTTCATCCCAACCCCTTA | 672 |
| QY | 1188 | gacttt-gatttcaagcatctctgtctcgttcccccataagagagcccggaattctgctg | 1248 |
| Db | 673 | GACTTTGGATTTTACGCAATCTTGCTTCTGTCTCCCATGAGACTCCGAATCTTGCGTG | 732 |
| QY | 1247 | tgtgtagatgtaggg 1260 | |
| Db | 733 | TGTGTAGATGAGGG 746 | |

RESULT 3

Q59349/c

Q59349 standard: cDNA: 430 BP.

Q59349

AC 059349:

DT 16-MAR-1994 (first entry)

DE Human brain Expressed Sequence Tag EST00378.

KW Gene transcription product; genetic markers; tagging; in vivo;

RW transcription; mapping; locations; chromosomes; chromosomal; ss.

OS Homo sapiens.

PN M0931617g-A.

PD 19-AUG-1993.

PR 12-FEB-1993; 001294.

PR 12-FEB-1992; US-831195.

PI (USSH) US DEPT HEALTH & HUMAN SERVICE.

PI Adams MD, Moreno RF, Venter CJ.

DR WPI: 93-272882/34.

PT Enriched oligonucleotides and corresp. sequences - used as

PT markers for human genes transcribed in-vivo, facilitate tagging

PT of most human genes

PS Example 4; Page 160: 500pp: English.

CC The Expressed Sequence Tag was isolated from a human brain cDNA

CC library as part of a large set of ESTs which can be used as markers

CC for human genes transcribed in vivo. They can be used to facilitate

CC tagging of most human genes, for mapping locations of expressed genes

CC on chromosomes, for individual or forensic identification, for mapping

CC locations of disease-associated genes, for identification of tissue

CC type, and for prepn. of antisense sequences, probes and constructs.

CC EST00378 has a "poor" coding probability as evaluated using the

CC coding-region prediction program CRM. See also Q59041-Q61440.

Q0 Sequence 430 BP; 109 A; 113 C; 105 G; 100 T;

| Query Match | 26.0% | Score 327.6 | DB 1 | Length 430 |
|-----------------------|---|--|--------------------|-------------|
| Best Local Similarity | 97.7% | Pred. No. 6.1e-90 | | |
| Matches 341 | Conservative | 0 | Mismatches 7 | Indels 1 |
| 913 | gggtcttcctcatcttaccac-aaggsggatacttgagtgcatcaattcccgagcaaggc | 971 | | |
| Db | 426 | gtttgcttccattttacacaaaggagatrttctgactgcatatattccccggacagagmc | 367 | |
| Qy | 972 | gaacttaaccctctccacatgysaaccttcctggggttggtgaacttgatgtgtta | 1031 | |
| Db | 366 | gaaacttaactctctcmccacatggaaccttctgggggttggaagaaactgattgtgtta | 307 | |
| Qy | 1032 | taaaagtgctccccaagcttggaagaccaggttggttaatactctggagacagccaaagc | 1091 | |
| Db | 306 | taaaaaagtgctccacagcttggaagaccaggttggttaatactctggagacagccaaagc | 247 | |
| Qy | 1092 | tgagatataaaggaggaaggaatgtagcaagaaacagagatcttcctgggttggtgtccc | 1151 | |
| Db | 246 | tgagatataaaggaggaaggaatgtagcaagaaacagagatcttcctgggttggtgtccc | 187 | |
| Qy | 1152 | cgctcccaactttccctttcatcccaacccccctagaacttgatttaacggatattctg | 1211 | |
| Db | 186 | cgctccctcaacttttccctttcatcccaacccccctagaacttgatttaacggatattctg | 127 | |
| Qy | 1212 | ctctgttcccatgagagctccgaattcttggtgtgtgtaagaggg | 1260 | |
| Db | 126 | cttctgttcccatgagagctccgaantcttggtgtgtgtaagatgaggg | 78 | |
| RESULT | 4 | | | |
| 059350 | | | | |
| ID | 059350 standard; cDNA; 317 BP. | | | |
| AC | 059350; | | | |
| DT | 16-MAR-1994 (first entry) | | | |
| DE | Human brain Expressed Sequence Tag EST00379. | | | |
| KW | Gene transcription product; genetic markers; tagging; in vivo; transcription; mapping; locations; chromosomes; chromosomal; ss. | | | |
| OS | Homo sapiens. | | | |
| PN | WO9316178-A. | | | |
| PD | 19-AUG-1993. | | | |
| PF | 12-FEB-1993; U01294. | | | |
| PR | 12-FEB-1992; US-837195. | | | |
| PA | (USSH) US DEPT HEALTH & HUMAN SERVICE. | | | |
| PI | Adams MD, Moreno RF, Venter CJ; | | | |
| PR | Enriched oligonucleotides and corresp. sequences - used as | | | |
| PT | markers for human genes transcribed in-vivo, facilitate tagging | | | |
| PS | of most human genes. | | | |
| CC | Example 4; Page 160; 500pp; English. | | | |
| CC | The Expressed Sequence Tag was isolated from a human brain cDNA | | | |
| CC | library as part of a large set of ESTs which can be used as markers | | | |
| CC | for human genes transcribed in vivo. They can be used to facilitate | | | |
| CC | tagging of most human genes, for mapping locations of expressed genes | | | |
| CC | on chromosomes, for individual or forensic identification, for mapping | | | |
| CC | locations of disease-associated genes, for identification of tissue | | | |
| CC | type, and for prep. of antisense sequences, probes and constructs. | | | |
| CC | EST00379 has a "poor" coding probability as evaluated using the | | | |
| CC | coding-region prediction program CRM. See also Q59041-Q61440. | | | |
| CC | Sequence 317 BP; 79 A; 71 C; 110 G; 54 T; | | | |
| Qy | Query Match | 7.3% | Score 91.6 | DB 1 |
| Db | Best Local Similarity | 81.5% | Pred. No. 2.1e-18; | Length 317; |
| Matches | 106; Conservative | 0 | Mismatches 24; | Indels 0; |
| Qy | 502 | tatccctgacagagctccgcgagacagatctccatgcccctggaaagcctggaaatgg | 561 | |
| Db | 149 | TAACCTTACACCTCTTTCCATGACAGAGATTCCTGATGCCCTTGAGAGCTGG | 208 | |
| Qy | 562 | gagagatcccgaaaggagagacagtgtgtcaccacaaacagaaagacagacactgttc | 621 | |
| Db | 209 | GAGAGATCCCGAAAGGAGAGACAGTGTGTCAACCCAAACAGAAAGAGTGAAGGCTTC | 268 | |

KW Batten disease; ceroid lipofuscinosis; CLN3; diagnosis;
KM gene therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 138..839
FT mutation /tag= a
FT mutation /note= "695delAG 2 bp deletion in Bd family L116"
PN MO9708308-A1.
PD 06-MAR-1997.
PE 31-AUG-1996; U13896.
PR 31-AUG-1995; US-003030.
PA (GEHO) GEN HOSPITAL CORP.
PA (UYLE-) RIKSUNIV LEIDEN.
PI Breuning MH, Gardiner MR, Gusella JF, Lerner TJ;
PI Mole SE, Taschner PEM;
DR WPI; 97-179265/16.
DR P-PSDB; W13585.
PT Batten disease polypeptide - useful to correct absence of wild type
PT polypeptide, or as agonist to enhance activity of wild type
PT polypeptide.
PS Disclosure; Page -: 94pp; English.
CC This sequence shows a novel mutation of the human CLN3 Batten
CC disease (Bd) gene, corresponding to a deletion of bases 695-696 of
CC CLN3 cDNA clone cDNA2-3 (T61306). 19 Novel mutations of the CLN3
CC gene were identified in 32 Bd patients. 17 of these mutations
CC (T61332-48) being in exon locations. The mutations were identified
CC by PCR amplification of CLN3 exons (see also T61331, T77158-86),
CC single strand conformation polymorphism analysis and direct
CC sequencing. The 695delAG mutation, in exon 8, results in a
CC frameshift after amino acid 5185 in the mutant CLN3 polypeptide
CC (W13585). It was found in a Bd family of Italian.
SQ Sequence 1730 BP; 316 A; 546 C; 464 G; 404 T;

Query Match 3.0%; Score 37.4; DB 1; Length 1730;
Best Local Similarity 55.9%; Pred. No. 0.14;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 436 cagagcctcagagagaggtgagccgctcagagagagagccctccccaagtgg 495
DB 630 CAAAGCCTGATGAGATCTACGCAAGACCAACACAGCGCTGTCCCAAGAAATGAG 571
QY 496 gaagggtatccctgcagagctcccgagagagaggttcgcatgcctggagagcctggag 555
DB 570 AAAAGGCAACGAGAGCAAGCTTCAGAGCAACAATCCACTGACGAGAACCGGGGGC 511
QY 556 aatgggg 562
DB 510 TGTAGCG 504

RESULT 15
T61341/C
ID T61341 standard; cDNA; 1732 BP.
AC T61341;
DT 24-OCT-1997 (first entry)
DE Batten disease gene mutation in family L285.
KM Batten disease; ceroid lipofuscinosis; CLN3; diagnosis;
KW gene therapy; ss.
OS Homo sapiens.
FH Key Location/Qualifiers
FT cds 138..1454
FT mutation /tag= a
FT mutation /tag= b
FT mutation /note= "G1020A missense mutation in Bd family
L285"
PN MO9708308-A1.
PD 06-MAR-1997.
PE 30-AUG-1996; U13896.
PR 31-AUG-1995; US-003030.

PA (GEHO) GEN HOSPITAL CORP.
PA (UYLE-) RIKSUNIV LEIDEN.
PI Breuning MH, Gardiner MR, Gusella JF, Lerner TJ;
PI Mole SE, Taschner PEM;
DR WPI; 97-179265/16.
DR P-PSDB; W13586.
PT Batten disease polypeptide - useful to correct absence of wild type
PT polypeptide, or as agonist to enhance activity of wild type
PT polypeptide.
PS Disclosure; Page -: 94pp; English.
CC This sequence shows a novel mutation of the human CLN3 Batten
CC disease (Bd) gene, corresponding to a G1020A nucleotide change in
CC CLN3 cDNA clone cDNA2-3 (T61306). 19 Novel mutations of the CLN3
CC gene were identified in 32 Bd patients. 17 of these mutations
CC (T61332-48) being in exon locations. The mutations were identified
CC by PCR amplification of CLN3 exons (see also T61331, T77158-86),
CC single strand conformation polymorphism analysis and direct
CC sequencing. The G1020A mutation, in exon 11, results in an E295K
CC amino acid change in the mutant CLN3 polypeptide (W13586). It was
CC found in a Bd family (L285) of Finnish origin.
SQ Sequence 1732 BP; 318 A; 546 C; 464 G; 404 T;

Query Match 3.0%; Score 37.4; DB 1; Length 1732;
Best Local Similarity 55.9%; Pred. No. 0.14;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

QY 436 cagagcctcagagagaggtgagccgctcagagagagagccctccccaagtgg 495
DB 630 CAAAGCCTGATGAGATCTACGCAAGACCAACACAGCGCTGTCCCAAGAAATGAG 571
QY 496 gaagggtatccctgcagagctcccgagagagaggttcgcatgcctggagagcctggag 555
DB 570 AAAAGGCAACGAGAGCAAGCTTCAGAGCAACAATCCACTGACGAGAACCGGGGGC 511
QY 556 aatgggg 562
DB 510 TGTAGCG 504

Search completed: May 14, 1999, 11:42:53
Job time: 5282 sec

GenCore version 4.5
Copyright (c) 1993 - 1998 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 13, 1999, 19:44:17 ; Search time 1016.5 Seconds
(without alignments)
1890.836 Million cell updates/sec

Title: US-09-212-270-10

Perfect score: 1260
Sequence: 1 cttcttagggagcctggaac.....tcgctgtgtgtatgtaggg 1260

Scoring table: IDENTITY_NUC

Searched: 2002476 seqs, 762712212 residues

Database :

EST.*
1: em_est1.*
2: em_est2.*
3: em_est3.*
4: em_est4.*
5: em_est5.*
6: em_est6.*
7: em_est7.*
8: em_est8.*
9: em_est9.*
10: gb_est1.*
11: gb_est2.*
12: gb_est3.*
13: gb_est4.*
14: gb_est5.*
15: gb_est6.*
16: gb_est7.*
17: gb_est8.*
18: gb_est9.*
19: gb_est10.*
20: gb_est11.*
21: gb_est12.*
22: gb_est13.*
23: gb_est14.*
24: gb_est15.*
25: gb_est16.*
26: gb_est17.*
27: gb_est18.*
28: gb_est19.*
29: gb_est20.*
30: gb_est21.*
31: gb_est22.*
32: em_est1.*
33: em_est2.*
34: em_est3.*
35: em_est4.*
36: em_est5.*
37: em_est6.*
38: em_est7.*
39: em_est8.*
40: em_est9.*
41: em_est10.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total score distribution, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB | ID | Description |
|------------|-------|-------------|--------|----|----------|--------------------|
| 1 | 481.4 | 38.2 | 495 | 24 | AA405973 | AA405973 zu57g07.r |
| 2 | 423 | 33.6 | 441 | 21 | AA477087 | AA477087 zu35f06.r |

| Result | Score | Query Match | Length | DB | ID | Description |
|--------|-------|-------------|--------|----|----------|--------------------|
| 1 | 481.4 | 38.2 | 495 | 24 | AA405973 | AA405973 zu57g07.r |
| 2 | 423 | 33.6 | 441 | 21 | AA477087 | AA477087 zu35f06.r |

ALIGNMENTS

RESULT 1
AA405973
LOCUS zu57g07.r1 Soares ovary tumor NBDOT Homo sapiens cdNA clone 742140
DEFINITION 5' mRNA sequence.
ACCESSION AA405973
NID 92064008
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 495)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Getzel, G., Jost, S.,
Kritman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,
Martin, J., Moore, B., Schellberg, K., Stepien, M., Tan, F.,
Theising, B., White, Y., Wylie, T., Waterston, R., and Wilson, R.
WashU-NCI Human EST Project
Unpublished (1997)

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810


```

|||||
DB 259 GGAACCTTCTCGGCTTTTGAACAGCTATGTCTATTAAAAAGGCGCTCCAGCTTGG 318
OY 1054 aagaccagggtggtgacacactgagacagcagaagctgagatataaaggagagga 1113
DB 319 AAGACCAGGAGTGGTACATCTAGAGACAGCCAGAGCTGAGATATATAAGAGAGAGGA 378
OY 1114 tgtgacagagacagagagcattccctggtgttggctcccgcttccctacttcccttcc 1173
DB 379 TGTGAGGAGCAAGAGCATCTTCTGTGGCTTCCCGTTCCTCATTTCCTTCCCTTTC 438
OY 1174 att 1176
DB 439 ATT 441

RESULT 3
AA291913 438 bp mRNA EST 08-AUG-1997
LOCUS zt38h12.r1 Soares ovary tumor NBHOT Homo sapiens CDNA clone 724679
DEFINITION 5', mRNA sequence.
ACCESSION AA291913
KEYWORDS 91939891
SOURCE EST.
ORGANISM human.
Eukaryote: mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
1 (bases 1 to 438)
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,
Parsons, J., Rifkin, B., Rohlfing, T., Tan, F., Trevaskis, E.,
Waterston, R., Williamson, A., Wohlmann, P., and Wilson, R.
WashU-Merck EST Project
unpublished (1995)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1277 Std Error: 0.00
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 426.
Location/Qualifiers
1. 438
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGAGCGAGCGCGCGGTGTTTATTTTATTTT 3']
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone="724679"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
<1..>438

MRNA
BASE COUNT 78 a 147 c 106 g 107 t
ORIGIN
Query Match 33.5%; Score 422.2; DB 21; Length 438;
Best Local Similarity 99.1%; Pred. No. 8.3e-110;
```

```

Matches 435: Conservative 0: Mismatches 3: Indels 1: Gaps 1:
OY 1 cttctagagaggagctgaaacctaattctctgagcctgagggaggggtgaggtcctcaag 60
DB 1 CTTCCTAGAGGAGCTGAACTAATTCTCTGAGCTGAGGAGGGGTGAGGGTCTCAAG 60
OY 61 gcaagctggtgcccacagagagagtgccagagagcactaaagtaaccttgccttcc 120
DB 61 GCAAGCTGGCCCCACAGAGAGGTGCCAGAGACACTAACAGTACCTTAGCTGTTC 120
OY 121 tctctccctcttatttcttaagttcccttatttcttccctggtgaacacctcttc 180
DB 121 TCTCTCCCTCTTTTATTTCAGTTCCTTTTATTTCCTTCCTGCTAACACTTCTTC 180
OY 181 cctctgcaacactgcccgtacaccttaaccgcccgcacacctctgtactaccaactct 240
DB 181 CCTCTGACACACTGCGCCCTTACCTTACCAGCCGCCGACCTCTGTGTAACCCACTCT 240
OY 241 gaaacacagctgttgcaaggttcccacagctatgccaagctcattctctctgtctag 300
DB 241 GAAACCAACAGCTGTGGCGAGGCTCCCACTCATGCTCATCTCTTCTTGTCTAG 300
OY 301 cccccaagagcctcagaggaacatgaggggccagatcagagagccggacctcagttgc 360
DB 301 CCCCAGAGGCTCCAGG-AACATGGGGGGCCAGTCAGAGAGCCGCACTCTCACTTGC 359
OY 361 cctctggtgagttgaggagcagctctgaggagcgtgtgttgcatgttgcagctgacc 420
DB 360 CCTCTGTTGAGTTGGGGGCGAGCTCTGGGGGCGCTGTGTCATGCTGTGCTGAC 419
OY 421 caacaacagagctgcaga 439
DB 420 CAACAACAGAGCTGCAGA 438

RESULT 4
AA481449 479 bp mRNA EST 08-AUG-1997
LOCUS zva5col.s1 Soares ovary tumor NBHOT Homo sapiens CDNA clone 756576
DEFINITION 3', mRNA sequence.
ACCESSION AA481449
KEYWORDS 92211001
SOURCE EST.
ORGANISM human.
Eukaryote: mitochondrial eukaryotes; Metazoa; Chordata;
Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Homnidae;
Homo.
1 (bases 1 to 479)
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Jost, S.,
Kucaba, T., Lacey, M., Le, N., Lennon, G., Marra, M., Martin, J.,
Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B.,
White, Y., Wylie, T., Waterston, R., and Wilson, R.
WashU-Merck EST Project 1997
unpublished (1997)

TITLE
JOURNAL
COMMENT

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m13 fwd. ET from Amersham
High quality sequence stop: 468.
Location/Qualifiers
1. 479
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGAGCGAGCGCGCGGTGTTTATTTTATTTT 3']
FEATURES
SOURCE
```

double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo.

/db_xref="taxon:9606"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>479)
BASE COUNT 122 a 128 c 121 g 108 t
ORIGIN

Query Match 33.4%; Score 421.4; DB 21; Length 479;
Best Local Similarity 99.8%; Pred. No. 1.4e-109;
Matches 422; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 838 cagagacatctatccgagtataaagagatccctccacggacgggacctaac 897
|||||
DB 479 CAGAGACATCTATCCCATATTAAGAGTAAGCCCTCCACCGGACGGGCTTACAC 420
OY 898 agctgctatagcgaagtgctctccattacacaaaggagatctgagtgacataat 957
|||||
DB 419 AGCTGCTATAGCGCAGGTGCTTCCATTTCACCAAGGGATATTCGAGTGCATATT 360
OY 958 ccccgagcaaggcggaacttaacctctccacatggaacctctcgggtttgtaaa 1017
|||||
DB 359 CCGCGGCAAGGGCGAAGCTTAACCTCTCCACATMAACCTCCGCGGAGTTGTGAAA 300
OY 1018 ctgtgattgtgtataaagtgctcccgacttggaagacgaaggtgtacatactg 1077
|||||
DB 299 CTGTGATGTGCTATATAAAGTGCTCCGACTTGGAAGACAGGGTGCTACTG 240
OY 1078 agacagccaagagctgagatataaagagaggaatgctgagaaagaggaacttcc 1137
|||||
DB 239 AGACAGCCAGAGCTGAGTATATAAGAGAGGGAATGTCAAGAAACAGAGGCACTCTCC 180
OY 1138 tgggtttggtcccgctctcaactttccctttcaatccacacccctgaacttgat 1197
|||||
DB 179 TGGTTTGGCTCCCGCTCCCTCACTTTTCCCTTTTCATTCACACCCCTGAGACTTGATT 120
OY 1198 ttccgatatctgctctctgtccccaagagctccgaattcttgctggttataagta 1257
|||||
DB 119 TTTCGATATCTGCTTCTCTTCCCATGAGCTCCGAATTCCTGCTGTGTAGATGA 60
OY 1258 ggg 1260
|||
DB 59 GGG 57

RESULT 5
AA443577/c 439 bp mRNA EST 03-JUN-1997
LOCUS 2w35801.s1 Soares ovary tumor NBHOT Homo sapiens cDNA clone 771240
DEFINITION 3', mRNA sequence.
ACCESSION AA443577
NID 92156252
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 439)
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J., Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theisling, B., White, T., Wylie, T., Waterston, R., and Wilson, R.
TITLE Washu-Merck EST Project 1997
JOURNAL Unpublished (1997)

COMMENT

Contact: Wilson RK
Washu-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -41m3 fwd. ET from Amersham
High quality sequence stop: 401.
Location/Qualifiers

FEATURES

source

1. 439
/organism="Homo sapiens"
/note="Organ: ovary; Vector: pT73D (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCACATCTGAGTGGAGCGGCGCGGTTTTTTTTTTT 3']".
double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo.
/db_xref="taxon:9606"
/clone_lib="Soares ovary tumor NBHOT"
/sex="Female"
/tissue_type="ovarian tumor"
/lab_host="DH10B (ampicillin resistant)"
complement(<1..>439)
/db_xref="GDB:598116"
BASE COUNT 87 a 127 c 116 g 109 t
ORIGIN

Query Match 32.5%; Score 409.8; DB 20; Length 439;
Best Local Similarity 99.3%; Pred. No. 2.8e-106;
Matches 422; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

OY 513 gaagtcgccgagacagagtcgagtcgcttggaagcctggaagatggaagatcccg 572
|||||
DB 439 GAGTCTCCGGAGACAGAGTTCGATCCCTCGAAGCCTGGAGAAATGGGAGATCCCG 380
OY 573 gaaagaagagcagctgtctaccccaaaacagaagaagcagcctgtctgcactggt 632
|||||
DB 379 GAAAGAGAGCAGTGTCTCACCCAAAACAGAGAGACACTGTCTGCACTGTGT 320
OY 633 ttccattacgcccactcccaagatgactccgattgacagagagtgatgtgcaaccagc 692
|||||
DB 319 TCCCATTTAAGCCACCTCCAGAGATACCTCCGATGTGACAGAGATGTGGCAACACAG 260
OY 693 tcttaggcgttggaagagcctacagcccaagatataggtgtccgaatccagatgctg 752
|||||
DB 259 TCTTAGGCGTGGGAGA-GCCTACAGGCCCAAGAGATGTGTGCCGAATCCAGATGCTGG 201
OY 753 agttatctgcgtgataagcagagctctgtttcaagaagtgactttcacatggtcagt 812
|||||
DB 200 AGTTTATCTGCTGTATAGGACAGTCTGTTCAGACGAGTTCACCATGGGTAGGT 141
OY 813 ggtgtctcagaagagcgaagagagagagactcattccgattgataagaagtatgac 872
|||||
DB 140 GGTGTCTCGAAGAGCCCAAGAGAGGAGAGACTTATTCCTCGATATAGAGATGTGTC 81
OY 873 ctccacccgagacgggactacacagctgctatagcaggtgtcttcattacacaca 932
|||||
DB 80 CTCACACCGGAGCGGCGCTACACAGCTGCTATACGAGGAGTGTTCATTACACT 21
OY 933 aagggt 937
|||||
DB 20 CGGGG 16

| | | | |
|-----------------|--|---|-------------|
| Db | 231 | CTACCAAGCTGCTATTAAGCGCAGGTGTCCTCATTTTACCACAAGGGGATATTCTGAGTGT | 172 |
| Oy | 951 | câtaattcccccgggaaggaggaacttaacctctcaccatggaacttccttggtt | 1010 |
| Db | 171 | CATAATTTCCCGGGGAAGGGCGAANCTTAACCTCTCTCCACTGGAACTTCTTG6G6FTT | 112 |
| Oy | 1011 | tgtgaacttgtatgtgtttaaaaaagtggctccccagcttggaaagacagagttgtgttac | 1070 |
| Db | 111 | TGTGAAGCTGATGTCTTTATTAANAAGTGGCTCCACACTTGGAAAGACAGGTGGTGCATC | 52 |
| Oy | 1071 | atactgsgaacagccaagagctgagtatataaagaagagaggaatgtgcagg | 1121 |
| Db | 51 | ATACGAGACAGCCAACAGCTGACTATATTAAGAAGAGAGGAATGTGCAGG | 1 |
| RESULT | 7 | | |
| LOCUS | AII149460/c | 462 bp mRNA EST | 28-OCT-1998 |
| DEFINITION | qc78b07.x1 Soares placenta_8to9weeks_2NDHPeto9w Homo sapiens cDNA | | |
| ACCESSION | AII149460 | | |
| NID | 93677929 | | |
| KEYWORDS | EST. | | |
| SOURCE | human. | | |
| ORGANISM | Homo sapiens | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria; | | |
| AUTHORS | Primates; Catarrhini; Homnidae; Homo. | | |
| TITLE | 1 (bases 1 to 462) | | |
| JOURNAL COMMENT | NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap. National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997) | | |
| FEATURES | Contact: Robert Strausberg, Ph.D. Tel: (301) 496-1550 Email: Robert_Strausberg@nih.gov This clone is available royalty-free through LNL ; contact the IMAGE Consortium (info@image.lnl.gov) for further information. Insert Length: 607 Std Error: 0.00 Seq primer: -40ml3 fwd. ET from Amersham High quality sequence stop: 457. Location/Qualifiers 1. 462 /organism="Homo sapiens" /note="Organ: placenta; Vector: pT73D (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5' TGTTCACATCTGAGTGGAGGCGGCCGCGCATTTTTTTTTTTTTTTT 3'] , double-stranded cDNA was size selected, ligated to Eco RI adapters (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT73 vector (Pharmacia). Library constructed by Bento Soares and M.Fátima Bonaído." /db_xref="taxon:9606" /clone="IMAGE:11715701" /clone_lib="Soares-Placenta_8to9weeks_2NDHPeto9W" /dev_stage="Two placenta: one from 8 weeks and another from 9 weeks post conception" /lab_host="DH10B (ampicillin resistant)" | | |
| BASE COUNT | 114 a 126 c 120 g 102 t | | |
| ORIGIN | | | |
| Query Match | 31.3%; Score 394.4; DB 31; Length 462; Best Local Similarity 99.7%; Pred. No. 6.7e-102; | | |
| Matches | 395; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | | |
| Oy | 865 agtatgccctccaccagcgagccggagctacaacagctctataagcagagtgcttcctcat | 924 | |
| Db | 462 ACTATGCCCTCCACCCGAGCCGGCCATNACAACACTCTATAGCGCACGTCCTTCCAT | 403 | |
| Oy | 925 ttacccaaggagatctcagttcgtatcataattccccggggaaggcgaaacttaactc | 984 | |

| | | | |
|----|------|---|------|
| Db | 402 | TTACACCAAGGGGATATTCTGAGTGCATTAATTCCTCCGGGCAAGGGCGAAACTTAACTC | 343 |
| QY | 985 | ttctccatcttggaacctctccgggtgttctgtaacctcttgatctgtgtataataaagtggctc | 1044 |
| Db | 342 | TTCTCAATGGAACCTTCCTGGGGTTTGTAACCTGTGATGTGTTATAAAAAGTGGCTC | 283 |
| QY | 1045 | ccagcttggagaagccaggctgggacacactctggagacagcccaagagctagatataaag | 1104 |
| Db | 282 | CCACCTTGGGAAGACACAGGTGTGGTACACTGTGAGACAGCACAAGAGCTGATATATAAG | 223 |
| QY | 1105 | gaagggaatgtctcaggaataaaggcatctctcttggtgttggctccccgttccctcaatt | 1164 |
| Db | 222 | GAGGGGAATGTGAGGAACGAGAGGCTCTCTCTGGGTTTGGCTCCCGCTTCTCAATT | 163 |
| QY | 1165 | tcacctttcatctccacacccctcagactcttgattcttaacgataactctgctctgttcccca | 1224 |
| Db | 162 | TCCCTTTTCATTCACACCCCTGAGACTTGAATTAACGATATACCTGTCCTGTCCCA | 103 |
| QY | 1225 | tggagctccgaattcttgcgtgtgtgtatagatgggg | 1260 |
| Db | 102 | TGGAGCTCCGAATTCCTGCTGTGTGTGAGATGAGGG | 67 |

| | RESULT | 8 |
|------------|--|----------------|
| AA056924 | | |
| LOCUS | AA056924 | 467 bp mRNA |
| DEFINITION | z7796-08.r1 Soares pregnant uterus NbHPV Homo sapiens CDNA clone | 11-MAY-1997 |
| ACCESSION | 489038.5 | mRNA sequence. |
| NID | AA056924 | |
| KEYWORDS | g1549582 | |
| SOURCE | EST. | |
| ORGANISM | human. | |
| | Homo sapiens | |

| REFERENCE | AUTHORS | TITLE | JOURNAL |
|--------------------|---|------------------------|--------------------|
| 1 (pages 1 to 467) | | | |
| | Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M., Holman, M., Hultman, M., Kuchala, T., Le, M., Lennon, G., Maitra, M., Parsons, J., Rifkin, L., Rohlfing, I., Tan, F., Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and Wilson, R. | Wash-Merck Est Project | Unpublished (1995) |

Contact: Wilson RK
Washu-Merck EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 266 1800
Fax: 314 266 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through IHLN; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 867 Std Error: 0.00
Seq primer: -20m13 rev2 from Amersham.

```

FEATURES
source
Location/Qualifiers
1..467
/organism="Homo sapiens"
/note="Organ: uterus; Vector: p773-Pac; Site_1: Not I
Site_2: Eco RI; 1st strand cDNA was primed with a Not
oligo(dn) primer 15'
AATGGAGAAGATTCGGGCGCCCTTTTTTTTTTTTTTTT 3' },
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the N
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by M. Fatima Bernaldo."
/db_xref="taxon:9606"
/clone="1489038"
/clone_1lb="Soares pregnant uterus NBHPU"
/sex="female"
/dev_stage="adult"
/lab_host="DH10B"

```

| MRNA | <1. | .>467 |
|------------|-------|------------------|
| BASE COUNT | 112 a | 113 c 149 g 84 t |
| ORIGIN | | 9 others |

| | | | | | | | |
|-------------|------------|--------------|-----------|------------|----|--------|-----|
| Query Match | 30.7% | Score | 387.4 | DB | 19 | Length | 467 |
| Best Local | Similarity | 95.0% | Pred. No. | 6.5e-100 | | | |
| Matches | 438 | Conservative | 0 | Mismatches | 19 | Indels | 4 |
| | | | | | | Gaps | 4 |

| | | | |
|----|-----|---|------|
| OY | 378 | ggcagctcgggggcccgggttggttggtgcattggttcacgtacatacccaacaagaagctgcga | 437 |
| Db | 6 | gggcacgannctggggccgtgattagtcgatgctttagctacccacacagacgttgcga | 65 |
| OY | 438 | gagcctcagagaaaggtgtgaacccgctgtgcaagggacagagagccctccagaatggga | 497 |
| Db | 66 | ga-nctcagagagagaggtgagcccgctgcaggggacagagagccctccagaat-ggga | 123 |
| OY | 498 | agggtatcccttggcagagttctcccgagcagagtccgaagtcccttgaagccttggggaa | 557 |
| Db | 124 | agggtatcccttggcagagttctcccgagcagagtccgaagtcccttgaagccttggggaa | 183 |
| OY | 558 | tggggagagatcccggaagaagagagcagtgctaccaccaaaacagagaagcagcctc | 61.7 |
| Db | 184 | tggggagagatcccggaagaagaacagatgcttaccaccaaaacagagaagcagcctc | 24.3 |
| OY | 618 | tgtcttgacactgtgtcccatlaacgcacactccaaagatgactcga-tgtgaacaag | 67.6 |
| Db | 244 | tgtcttgacactgtgtcccatlaacgcacactccaaagatgactcga-tgtgaacaag | 30.3 |
| OY | 677 | tgatgtggcaaccagctctttaggcgttgggaagagcctacagagcccaagatagtctcc | 73.6 |
| Db | 304 | tgatgtggcaaccagctctttaggcgttgggaagagcctacagagcccaagatagtctcc | 36.2 |
| OY | 737 | gaatcagaagatgcttggagtttatctgtcgtataagcagagctcgttttcaagaagctgact | 79.6 |
| Db | 363 | gaatcagaagatgcttggagtttatctgtcgtataagcagagctcgttttcaagaagctgact | 42.2 |
| OY | 797 | tcaaccatggtcagggtggtgtcttcgaaagcgaagaag | 83.7 |
| Db | 423 | tcaccatgggtcagggtggtgtcttcgaaagcgaagaag | 46.3 |

| | | | | | | | | | |
|------------|--|--|--|--|--|--|--|--|--|
| RESULT | 9 | | | | | | | | |
| LOCUS | AA293679/c | | | | | | | | |
| DEFINITION | zC38H12.s1 Soares ovary tumor NBHOT Homo sapiens CDNA clone 724679 | | | | | | | | |
| ACCESSION | AA293679 | | | | | | | | |
| NID | G1941461 | | | | | | | | |
| KEYWORDS | EST. | | | | | | | | |
| SOURCE | human. | | | | | | | | |
| ORGANISM | Homo sapiens | | | | | | | | |
| REFERENCE | Eukaryotes; mitochondrial eukaryotes; Metazoa; Chordata; | | | | | | | | |
| AUTHORS | Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; | | | | | | | | |
| | Homom. | | | | | | | | |
| | 1 (bases 1 to 435) | | | | | | | | |
| | Hillier,L., Clark,N., Dubugue,T., Elliston,K., Hawkins,M., | | | | | | | | |
| | Hollerer,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., | | | | | | | | |
| | Peterson,J., Rifkin,L., Rohlfing,T., Tan,F., Treviskis,E., | | | | | | | | |
| | Matsenon,R., Williamson,A., Wohlmann,P. and Wilson,R. | | | | | | | | |
| TITLE | WASHU-Merck EST Project | | | | | | | | |
| JOURNAL | Unpublished (1995) | | | | | | | | |
| COMMENT | | | | | | | | | |

Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: estevatson.wustl.edu
This clone is available royalty-free through LINL ; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1277 Std Error: 0.00

Matches 382; Conservative 0; Mismatches 16; Indels 1; Gaps 1;

```
QY 169 aaaaactttctcctctcaccacacgcccgtacaccttaaccgccccgacctcttc 228
Db 1 ACAAACCTTTCTCCCTCTCACCACACTGCCGTACCTTACCCGCCGCCACCTCTTGC 60
QY 229 taaccacacttgaaccacacagctgttgcaagggtccacagctcagccacatctc 288
Db 61 TACCCACACTTGAACACACAGCTGTGGCAGGGGTCCCAAGCTCATGCCACTCATTC 120
QY 289 cttctctgtagcccccacaaaggccctccagcaacatggggggccagtcagagaccg 348
Db 121 CTTTCTTGTAGCCCCCAAGCCCTCCAGCAACATGGGGGCCCAAGTCAGAGACCCGG 180
QY 349 actctcagttgcccctctgtgttggtgggggagctctcgtggggccgtgtgtgcatg 408
Db 181 ACTCTAGTTGCCCTCTGTGTTGAGTTGGGGGAGCTCTGGGGGCCGTGCTGTGCATG 240
QY 409 gtccagctaccacaacacagagctgcagagcctcagagagaaggtgaagcgtgcag 468
Db 241 CTNGAACCTACCCACACACAGAGCTGCAGACCTCAAGAGAGAGTGAACCCGCTGCAG 300
QY 469 gggacagagagccctccacgaatggggaaggatccctgagagctcccgagagca 528
Db 301 GGGACAGAGAG-CCTNCCAGATGANNAGGTATCCCTGNAGAGTCTCCCGAGAGCA 359
QY 529 agtccgagtcctcggagagcctgggaggaatgggagaga 567
Db 360 AGTCCGATGCCCTGGAGAGCTGGGAGAAATGGGGAGAGA 398
```

RESULT 13

AA790394 583 bp mRNA EST 06-FEB-1998
LOCUS vw04a1.r1 Soares mouse mammary gland NbMNG Mus musculus cDNA clone
DEFINITION 1230908 5', mRNA sequence.
ACCESSION AA790394
NID 92850514
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 583)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterston,R.

TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT

Contact: Marra M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.wustl.edu
This clone is available royalty-free through LML; contact the IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:656500
Putative full length read
vector to vector length is 891
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 499.

FEATURES

source 1..583
/organism="Mus musculus"
/strain="C57BL/6J"
/note="Organ: mammary gland; Vector: pT73D-Pac (pharmacia) with a modified polylinker. Site_1: Not I; Site_2: Eco RI; 1st strand cDNA was primed with a Not I -

oligo(dT) primer [5'
TGTTCACATCTGAAGTGGAGGCGCGCGGATGGCTTTTTTTTTTTTTTTTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
RNA provided by Dr. Minoru Ko, Wayne State Univ. Library
constructed and normalized by Bento Soares and M.Fatima
Bonaldi.
/db_xref="taxon:10090"
/clone_lib="Soares mouse mammary gland NbMNG"
/clone_1230908
/sex="male"
/tissue_type="mammary gland"
/dev_stage="4 weeks"
/lab_host="DH10B"
BASE COUNT 140 a 133 c 187 g 122 t 1 others
ORIGIN

Query Match 28 6%; Score 360.2; DB 26; Length 583;
Best Local Similarity 79.5%; Pred. No. 3.6e-92;
Matches 464; Conservative 0; Mismatches 114; Indels 6; Gaps 3;

```
QY 308 aggcctccagagcaacatggggggccacagcagagacggaactcagttccctcgg 367
Db 5 AAGCTCCAGGCCACATGGGGGGCTCAGTCAGAGAGCCAGCCCTTTCGTTGCTTTGG 64
QY 368 ttgagttggggggcagctcgtggggccgtgtgtgc-atggttcagctgaaccacaa 426
Db 65 TTGAGTTGGGGGCACTTCTGGGGGCTGTGACTGTGTGCTGCACACTACTGATCCAAAC 124
QY 427 acagagctcagagcctcagagagaggtgagccggtcagagggaacagagccctcc 486
Db 125 ACAGAGCTCCAAACCTAAGGCGAGGTGAGCGGCTCAGCGGAGTGGAGGCTTCC 184
QY 487 cagaaatggggaagggtatccctcgtgagagctcccgagacaaagtccgagctggaa 546
Db 185 CAGAGCAGAGGAGAGGCCCAATG--CAGAGCTTGGAGAGAGTCTGTGTCCTGGAA 242
QY 547 gacctggagaatgcccgaaagagagagagcagctctcaccccaaaacagaag 606
Db 243 GCCTGGAAGATGGGGCGAATCTCGAGAGAGAGCAAGTACTACCCAGAGCAAG 302
QY 607 aagcagcactcgtcctcagcctgtgtccattaacgcacctcacaagatgactccg 666
Db 303 AAGAAACACTCACTCCTGCATCTTTCAGATTAACTTACCTCCAAAG--GACTCTGAC 359
QY 667 gtagcagaaggtatgtggaacacagctcttagcgctgggaaggagcctaacagggc 726
Db 360 GTGACAGAGGTATGTGGCAACCACTTAAAGCTGGGAGAGCCTGGAGGCCCAAG 419
QY 727 tatggttcgaatccagagatgtgagttatctgtatagcagagtcctgttcaa 786
Db 420 GACATGTGAGAGTCTGGGACACTGGAATTATCTCTATATAGAGTCCCTTTCA 479
QY 787 gacgtgacttcaaccatgggtcaggtgtgtctcgagaagggccaaggaagcagag 846
Db 480 GATGTGACTTTCACATGAGGTGAGTGTATCTCGGGAAGAGCAAGGAGAAAGAACT 539
QY 847 ctattccgattgtaagaagtatgacctccacccgagcaggagc 890
Db 540 CTATTCGATGATACAGAGTATGCTTGTGATCCTGACCGTGC 583
```

RESULT 14

AI271780/c 473 bp mRNA EST 17-NOV-1998
LOCUS AI271780
DEFINITION qj82f03.x1 NCI_CGAP_K1d3 Homo sapiens cDNA clone IMAGE:1865981 3',
ACCESSION qj82f03
NID 93890947
KEYWORDS EST
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 473)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: M. Bento Soares, Ph.D.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 440.
Location/Qualifiers
1. 473
/organism="Homo sapiens"
/note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I oligo(dT) primer,
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT7T3 vector. mRNA
source: 2 pooled kidneys. Library went through one round
of normalization. Library constructed by Bento Soares and
M. Fatima Bonaldo."
/db_xref="taxon:9606"
/clone_image="1865981"
/clone_lib="NCI-CGAP_K1d3"
/lab_host="DH10B"
BASE COUNT 120 a 130 c 119 g 103 t 1 others
ORIGIN

Query Match 28.3%; Score 357; DB 31; Length 473;
Best Local Similarity 97.4%; Pred. No. 2.8e-91;
Matches 405; Conservative 0; Mismatches 6; Indels 5; Gaps 4;

845 cctcattccgagtcataagaagatgcccctccacccgagccgagctacacagctgct 904
|||||
467 cttctattccgagtcattat-agaagtatg-cctccacccgagccgagccct-CAACAGCTGCT 411
|||||
905 atagcgcagagtgcttcattacacacagggagatattctgagtgataattcccgagg 964
|||||
410 ATACGCGAGGtGtGTCCTTCACTTTA--ACCAAGGGGATATTCTAGTCAATAATTCCCGGG 353
|||||
965 caagggcgaacttaacctctccacatggaacctctcctgggggtttgtgaacctgtgat 1024
352 CAAGGGCGAAACTTAACCTCTCTCCACATGGAACCTTCTGGGTTGTGAAATGTGAT 293
|||||
1025 tctgttataaagtgctcccaagcttggaagcagaggtggtgatacactggaagcagc 1084
292 TGTGTTATAAAGTGTGCTCCACCTTGGAAGACGAGGTGGTATACATGAGAGACAGC 233
|||||
1085 caaagagctgaatatataaagagggagatgtgacgaagacagagcatcttcctgggttt 1144
232 CAAGAAGCTGAGTATATAAAGGAGGAGATGTGAGGAACAGAGGCGCTTCTCGGGTTT 173
|||||
1145 ggtcccccgttcctcacttccctttcattccacacccctagactgtgatttaacgga 1204
172 GGGTCCCGGTTCCACACTTTCCTTTTCATTTCCACCCCTAGACTTGTGATTTTACGGA 113
1205 taccctcttcgttccacatgagagctcgaattcttgcgtgtgtatgataagag 1260
|||||

Db 112 TACTTCTCTGTTCCCATGAGAGCTCCGATTTCTTGCTGTGTGATGAGAG 57

RESULT 15
AI281823/c 448 bp mRNA EST 23-NOV-1998
LOCUS q187a03.x1 NCI-CGAP_Col4 Homo sapiens cDNA clone IMAGE:1962220 3',
DEFINITION mRNA sequence.
ACCESSION AI281823
NID 93920056
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Mammalia; Eutheria;
Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 448)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT

Contact: Robert Strausberg, Ph.D.
Tel: (301) 496-1550
Email: Robert_Strausberg@nih.gov
Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdrp/image/image.html

Seq primer: -40UP from Gibco
High quality sequence stop: 430.
Location/Qualifiers
1. 448
/organism="Homo sapiens"
/note="Organ: COLON; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.7 kb. Life Technologies catalog #:
11531-019"
/db_xref="taxon:9606"
/clone_image="1962220"
/clone_lib="NCI-CGAP_Col4"
/tissue_type="moderately-differentiated adenocarcinoma"
/lab_host="DH10B"
BASE COUNT 113 a 123 c 114 g 97 t 1 others
ORIGIN

Query Match 27.7%; Score 349.4; DB 23; Length 448;
Best Local Similarity 99.2%; Pred. No. 4e-89;
Matches 361; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

897 cagctgtctaagcgcagtgcttcattacccaagggagatattctggttcaat 956
|||||
448 CAGCTGCTATAGCGCAGAGTGTCTTCACTTACACCAAGGGGATATCTGAGTGCATAAT 389
|||||
957 tccccggcgaaggcgaacttaacctctctccacatggaaccttcctgggttgaa 1016
448 CAGCTGCTATAGCGCAGAGTGTCTTCACTTACACCAAGGGGATATCTGAGTGCATAAT 389
|||||
388 TONCGGGCAAGGGCGCAAACTT-ACCTCTCTCCACATGGAACCTTCTGGGTTTGGA 330
|||||
1017 acgtgattgtgtataaagtgctccagcttggaagcagaggtggtacatactg 1076
329 ACTGTGATTGTGTATATAAAGTGTGCTCCACCTTGGAAGACGAGGTGGTATACATG 270
|||||
1077 gagcagcgaagcgtgagatataaagggagggatgagcaggaagcagcatcttc 1136
269 GAGACAGCCCAAGAGCTGATATATAAGGAGAGGAGATGTGCAAGAACAGAGCGCTTTC 210
1137 ctgggttcgctcccgcttcctcacttctcccttccatcccccagagacttgat 1196
|||||

Db 209 CTGGTTGGCTCCCGTTCCCTCACTTTCCCTTTTCATTCACCACCCCTAGACTTTGAT 150
QY 1197 tttaaggatactcttgcctctctgttccccaagagctccgaattcttgcgtgtgtatgacatg 1256
Db 149 TTACGGATATCTTCTCTCTCTCCCATGAGCTCCGATTCCTTGCCTGTGTGTAGATG 90
QY 1257 aggg 1260
Db 89 AGCG 86

Search completed: May 13, 1999, 19:44:22
Job time: 11327 sec

This Page Blank (uspio)